Metabolic Engineering and Molecular Tool Development in Yeast for Production of Bulk Chemicals

Fabre, Mathew Malcolm Jessop

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Metabolic Engineering and Molecular Tool Development in Yeast for Production of Bulk Chemicals

Ph.D. Thesis

Mathew Malcolm Jessop-Fabre

Novo Nordisk Foundation Centre for Biosustainability
Technical University of Denmark

Metabolic Engineering and Molecular Tool Development in Yeast for Production of Bulk Chemicals

Novo Nordisk Foundation Centre for Biosustainability

Technical University of Denmark

Cover Design: Representation of stained yeast cells by Shelby Criswell

novo nordisk fonden

DTU Biosustain
The Novo Nordisk Foundation Center for Biosustainability

Preface

This dissertation serves as partial fulfilment of the requirements to obtain a Ph.D. at the Technical University of Denmark, under the Centre for Biosustainability Ph.D. school. The work presented in this thesis was carried out at the Novo Nordisk Foundation Centre for Biosustainability, and partly at the Keasling Lab at the Joint BioEnergy Institute, University of California, Berkeley, United States of America between 2014 and 2018. This work was carried out under the supervision of Dr. Irina Borodina, DTU, and co-supervision of Dr. Yun Chen, Systems and Synthetic Biology group, Department of Biology and Biological Engineering, Chalmers University of Technology, Sweden. The work presented here has the aim of developing molecular tools for the metabolic engineering of *Saccharomyces cerevisiae*, and the application of these tools to investigate acid production in the host organism. This project was funded by the Novo Nordisk Foundation.

Mathew Malcolm Jessop-Fabre

August 2018

Abstract

The brewers' yeast *Saccharomyces cerevisiae* has long been a close friend of humanity. Yeast has been used in food and beverage production for thousands of years, and our close relationship with yeast has led to it becoming a model organism in research. Yeast was the first eukaryote to have its genome sequenced, and has helped in the study of human diseases, such as cancer. This extensive characterisation, along with its routine use in bioprocesses makes it an attractive target for metabolic engineering efforts.

The work presented in this thesis encompasses some of the work that has been undertaken as part of the Ph.D. programme. Herein, I present work towards the development of tools to improve the process of engineering *S. cerevisiae*, as well as to apply those and other existing metabolic engineering tools to gain insights into the metabolism of engineered yeast strains. This research focuses on the production of hydroxy acids, a class of carboxylic acid.

In the first research study, we developed a new set of integrative yeast expression vectors based on the previously developed EasyClone system. We adapted this system for use with CRISPR-Cas9 to insert expression cassettes into the yeast genome without the need for also integrating markers. This increases the speed of the engineering cycle as no marker loop-out is required between transformations. It also negates the need for markers, both auxotrophic and dominant, which can alter native host physiology and potentially affect experimental results. We have made this vector toolkit available through the Addgene platform, and hope many will use it to facilitate their research in *S. cerevisiae*. Using our vector toolkit, we investigated different acetyl-CoA supply strategies and their impact on the production of 3-hydroxypropionic acid (3HP), a bulk chemical that is a common target for metabolic engineering. We discovered that expression of a bacterial pyruvate

dehydrogenase complex was the most successful strategy, which increased 3HP titres by almost 100% when grown in a simulated fed-batch medium.

In the second study, focus is shifted away from tool development and towards metabolic engineering. Here we engineered a Crabtree-negative strain of *S. cerevisiae* to produce three different hydroxy acids; lactate, malate, and 3HP. We characterised these strains, and sought to gain an understanding of how these heterologous pathways influence the metabolism and physiology of the host strain. We performed ^{13}C-based metabolic flux analysis to quantify the carbon flux distributions through the central carbon metabolism. Between the strains, we identified large differences in flux distribution, but each strain showed surprisingly high levels of flux through the pentose phosphate pathway, and only low levels of flux through the tricarboxylic acid (TCA) cycle.

The third study attempts to use metabolic flux and transcriptomic analysis to identify key gene deletion strategies that can boost the production of 3HP in a host strain (ST938) capable of producing high levels of this hydroxy acid. Transcriptomic analysis revealed that much of central carbon metabolism was upregulated in ST938 compared to the WT strain, but metabolic flux analysis revealed lower proportional carbon flux through the TCA cycle. Metabolic flux analysis also revealed that only a small proportion of the available carbon flux was diverted into 3HP production. Computational approaches and insights from the transcriptomic analysis identified several gene deletion strategies that were then implemented *in vivo*. Of these strategies, the deletion of *PRY1* was able to increase final 3HP titres by 27%.

Dansk Resumé

Bryggergæren *Saccharomyces cerevisiae* har længe været en nær ven af menneskeheden. Gær har været brugt i mad og drikkevareproduktion i tusindvis af år, og vores tætte forhold til gær har ført til brugen af denne som modelorganisme i forskning. Gær var den første eukaryot der fik sit genom sekventeret og har været vigtig for forskning af menneskebårne sygdomme, såsom kræft. Den omfattende karakterisering i kombination med den rutinemæssige anvendelse i bioprocesser gør det til et attraktivt mål for metabolisk genmodificering.

Denne afhandling omfatter noget af det arbejde, der er udført som led i mine Ph.D studier. Heri præsenterer jeg arbejde i forbindelse med udvikling af værktøjer til forbedring af processen omkring genmodificering af *S. cerevisiae*, samt at anvende disse nyudviklede værktøjer sammen med allerede eksisterende værktøjer til at få indblik i metabolismen hos de nyudviklede gærstammer. Denne forskning fokuserer på produktion af hydroxysyrer, en klasse af carboxylsyre.

I det første studie udviklede vi et nyt sæt integrerede gærekspressionsvektorer baseret på det tidligere udviklede EasyClone-system. Vi tilpassede dette system til brug med CRISPR-Cas9 til at indsætte ekspressionskassetter i gærgenomet uden behov for også at integrere markører. Dette øger hastigheden, da der ikke kræves en markørsløjfe mellem transformationer. Det negerer også behovet for markører, både auxotrofe og dominerende, hvilket kan ændre indfødt værtsfysiologi og potentielt påvirke eksperimentelle resultater. Vi har gjort denne vektor værktøjskasse tilgængelig via Addgene platformen, og håber mange vil bruge den til at gøre deres forskning i *S. cerevisiae* nemmere. Ved hjælp af vores vektor værktøjssæt undersøgte vi forskellige acetyl-CoA forsyningsstrategier og deres indvirkning på produktionen af 3-hydroxypropionsyre (3HP), et kemikalie, der ofte bliver brugt som mål for metabolisk genmodificering . Vi fandt, at

ekspression af et bakterielt pyruvat-dehydrogenasekompleks var den mest vellykkede strategi, der forøgede 3HP-titrer med næsten 100%, når de blev dyrket i et simuleret fodret batch medie.

I det andet studie er fokus rykket væk fra værktøjsudvikling, og fokus er i stedet på metabolisk genmodificering. I dette studie konstruerede vi en Crabtree-negativ stamme af *S. cerevisiae* til fremstilling af tre forskellige hydroxysyrer; lactat, malat og 3HP. Vi karakteriserede disse stammer og søgte at få en forståelse for, hvordan disse heterologe veje påvirker værtsstammens metabolisme og fysiologi. Vi udførte ^{13}C-baseret metabolisk fluxanalyse til at kvantificere carbonfluxfordelingen gennem den centrale carbonmetabolisme. Vi identificerede store forskelle i fluxfordelingen mellem stammerne, men alle stammer viste overraskende høje niveauer af flux gennem pentosefosfatvejen og kun lave niveauer af flux gennem tricarboxylsyre (TCA) -cyklussen.

I det tredje studie forsøger vi at anvende metabolisk flux og transkriptomisk analyse til at identificere nøglestrategier for målrettet genfjernelse, der kan øge produktionen af 3HP i en værtsstamme (ST938), der allerede er kendt for at producere høje niveauer af denne hydroxysyre. Transcriptomisk analyse viste, at meget af den centrale carbonmetabolisme er opreguleret i ST938 sammenlignet med WT-stammen, mens metabolisk fluxanalyse afslørede at en lavere proportion af kulstoffet føres gennem TCA-cyklusen. Metabolisk fluxanalyse afslørede dog også, at kun en lille del af den tilgængelige carbonflux blev omledt til 3HP-produktion. Computerberegninger og indsigt fra den transkriptomiske analyse identificerede adskillige genfjernelsesstrategier, som derefter blev implementeret in vivo. Af disse strategier var sletningen af *PRY1* i stand til at øge de endelige 3HP titre med 27%.

Acknowledgements

I have learnt many new things during my time at CfB, both in and outwith science. Having not been exposed to a lab setting before starting the Ph.D., there was a steep hill to climb. All of my thanks and gratitude go to those who have helped me gain the skills that I need to undertake the work I have done. There are too many to thank individually, but all past and present members of both Yeast Metabolic Engineering and Synthetic Biology Tools for Yeast have not only helped me to learn, but also have provided a kind and supportive atmosphere to work in. My office mates too, for providing some very interesting discussions. Also important to mention are my new colleagues in the (formerly known as) iLoop unit who have been very supportive, with particular thanks to Niko Sonnenschein, and my office mates; Ali Kaafarani, Svetlana Kutuzova, Alba Lopez, Emre Özdemir, Moritz Beber, Pannipa Ponpitakpong, Alex Hoffmeyer, and Ibrahim El-Semman. There are also many people and friends that I have met during my PhD at CfB and DTU, unfortunately too many to mention all by name. But I wish to extend thanks to them all.

I could not have completed my studies without the supportive and patient hand of my supervisor, Dr. Irina Borodina without whom I would not have had this opportunity. Regards must also be paid to Dr. Itay Budin, with whom I worked in the Keasling Lab in the University of California, Berkeley.

Particular, individual, thanks need to be extended to the following; Jonathan, who provided me with many favours towards the end of my studies. Gheorghe, for his invaluable practical tutelage in the commencement of my PhD, and for many enjoyable discussions. Christian, for offering a rare and much appreciated kind of friendship, spurring a different kind of creativity than that found during my studies. Sarah, for always being so much fun, and for having the most interesting life of anyone that I know and managing to do all the

things I would never be able to. And lastly as always, Mikkel, who provides a much needed distraction from academic thought.

My final and most important thanks must be given to my family for all their support and to my girlfriend, Kathrine, for always being there for me. I cannot express my gratitude enough, but I hope to keep trying.

Contents

List of Publications

Paper I.

Mathew M Jessop-Fabre*, Tadas Jakočiūnas*, Vratislav Stovicek, Zongjie Dai, Michael K Jensen, Jay D Keasling, Irina Borodina. (2016) **EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into *Saccharomyces cerevisiae* via CRISPR-Cas9.** Biotechnology Journal. 11(8): 1110-1117.

Paper II.

Mathew M Jessop-Fabre, Jonathan Dahlin, Vratislav Stovicek, Itay Budin, Jay D Keasling, Irina Borodina. (2018) **A comparative study into the effects of different pyruvate derived product production pathways on the physiology and metabolism of a Crabtree-negative *Saccharomyces cerevisiae* strain.** Manuscript in preparation.

Paper III.

Mathew M Jessop-Fabre, Kanchana R Kildegaard, Birgitta E Ebert, Emre Özdemir, Lars M Blank, Irina Borodina (2018) **Omic data based improvement of a high performance 3-hydroxypropionic acid producing strain of *Saccharomyces cerevisiae*.** Manuscript in preparation.

List of Abbreviations

0-9

3HP	3-hydroxypropionic acid
3HPA	3-hydroxypropionaldehyde

A

ACL	ATP-dependent citrate lyase
ACS	Acetyl-CoA synthase
ADH	Alcohol dehydrogenase
AHA	α-hydroxy acid
ATP	Adenosine triphosphate

B

BAPAT	β-alanine-pyruvate amino transferase
BHA	β-hydroxy acid

C

CAD	Computer aided design
Cameo	Computer aided metabolic engineering and optimisation
Cas	CRISPR-associated system
cDNA	Complementary DNA
CDW	Cell dry weight
CoA	Co-Enzyme A
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CSM	Complete supplement mixture
CTR	Carbon transfer rate
C_2	Compound containing two carbon atoms

D

DHAP	Dihydroxy-acetone-phosphate
DNA	Deoxyribonucleic acid
DSB	Double-strand break

dsDNA	Double-stranded DNA

F

FBA	Flux balance analysis
FIT	Feed-in-time

G

GC-MS	Gas chromatography - mass spectrometry
GEM	Genome scale metabolic model
GFP	Green fluorescent protein
GO	Gene ontology
GRAS	Generally regarded as safe
gRNA	Guide RNA

H

HIBADH	3-hydroxyisobutyrate dehydrogenase
HPDH	3-hydroxypropionate dehydrogenase
HPLC	High-performance liquid chromatography
HR	Homologous recombination

I

INCA	Isotopomer network compartment analysis

L

LB	Lysogeny broth
LDH	Lactate dehydrogenase

M

MATLAB	Matrix laboratory
MFA	Metabolic flux analysis
MM	Minimal media
mRNA	Messenger RNA

N

NAD^+	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NGS	Next generation sequencing
NHEJ	Non-homologous end joining
NMR	Nuclear magnetic resonance

O

OD	Optical density
OHA	Ω-hydroxy acid
OTR	Oxygen transfer rate

P

PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PDC	Pyruvate decarboxylase
PDH	Pyruvate dehydrogenase
pFBA	Parsimonious flux balance analysis
pH	Potential of hydrogen
PP	Pentose phosphate

Q

qPCR	Quantitative PCR

R

RNA	Ribonucleic acid
RPM	Revolutions per minute
RQ	Respiratory quotient
rRNA	Ribosomal RNA

S

SBOL	Synthetic biology open language
SC	Synthetic complete
SD	Standard deviation

SDM	Site directed mutagenesis
sgRNA	Single guide RNA
SNP	Single nucleotide polymorphism

T

TALEN	Transcription activator-like effector nuclease
TCA	Tricarboxylic acid
tracrRNA	Transactivating CRISPR RNA
tRNA	Transfer RNA

U

USER	Uracil-specific excision reagent
UV	Ultra-violet

W

WT	Wild-type

Y

YPD	Yeast extract peptone dextrose

Z

ZFN	Zinc finger nuclease

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Thesis Overview

The yeast, *Saccharomyces cerevisiae*, is an important organism in research and in its use in many bio-industrial processes. Much work has been, and will continue to be, applied to the construction of yeast as a cell factory. The disciplines of Metabolic Engineering, Systems Biology, and Synthetic Biology are all integral parts in furthering the development of yeast and in helping to realise its potential as a source of biological and sustainable products.

This thesis offers a small addition to the field of yeast metabolic engineering, by contributing tools to engineer, and basic research into, yeast metabolism. The research contained herein focuses on the production of hydroxy acids in the model yeast organism *Saccharomyces cerevisiae*. Often talked about in metabolic engineering is the "Design-Build-Test-Learn" cycle, and its indispensability for progress within the field. Here, I present work that touches on all four of these segments, but with much of the focus on the testing side.

Chapter 1 introduces the themes that have surrounded the PhD studies and are foundational to the work undertaken therein. This chapter is not meant to serve as an exhaustive introduction to metabolic engineering in *S. cerevisiae*, as such an effort would fill an entire book. Instead, the topics related to the research work presented in the thesis are selected, in the aim of introducing the reader to the key topics of the thesis.

Chapter 2 is the first research article presented in the thesis, and focuses on the development of a CRISPR-Cas9 based toolkit for gene addition in the chromosomes of *S. cerevisiae* without the need to simultaneously insert marker cassettes. This system, named EasyClone-MarkerFree, builds upon the previous EasyClone vector sets and offers a method to increase the standardisation and reproducibility within the field through the use

of defined integration sites. We test this system in a diploid strain of *S. cerevisiae*, showing that it is possible to successfully use these vectors to insert genes into industrial strains, inserting gene copies into both chromosomes. We hope that people will use these vectors, along with the other EasyClone vectors, to facilitate their yeast genome engineering projects and reduce some of the complexity inherent in strain design. We demonstrate the practicality of this vector set by using it to test the impact of different cytosolic acetyl-CoA supply strategies on the production of the hydroxy acid, 3-hydroxypropionic acid (3HP). We discover that the functional expression of the pyruvate dehydrogenase complex from *Enterococcus faecalis* is a promising method for increasing acetyl-CoA pools in the cytosol.

Chapter 3 keeps with hydroxy acids, this time presenting a study on lactate and malate production as well as 3HP. In this study, we investigate the impact that these hydroxy acid production pathways have on the metabolism of a pyruvate over-producing Crabtree-negative yeast strain. This Crabtree-negative strain does not produce ethanol, making it an attractive host for the production of a wide range of chemicals that can be produced from pyruvate. In this study we use a mixture of different techniques, including ^{13}C-based flux analysis to probe the metabolic responses to high levels of lactate, malate, and 3-hydroxypropionic acid production. We show that all strains assimilate glucose primarily through the pentose phosphate pathway and have low fluxes through the TCA cycle. The strains show variation in lower glycolysis, with the 3HP strains showing high activity in glycerol-3-phosphate biosynthesis pathways, whereas the lactate and malate producers show reduced activity in these pathways.

The final article, presented in Chapter 4, utilises -omic scale technologies, this time with the aim of characterising a yeast strain capable of high levels of 3-hydroxypropionic acid production. Here we show that this high producer upregulates glycolysis, and components of the pentose phosphate pathway. ^{13}C-based flux analysis determined that the engineered

strain had much lower flux through the tricarboxylic acid cycle than the WT, due to pyruvate being drawn from this pathway into the synthesis of 3HP. However, we also show that not all of the available pyruvate is successfully converted to 3HP, revealing that there are still opportunities for further engineering of this biosynthetic pathway. We used a flux constrained genome scale metabolic model and the OptKnock algorithm to computationally predict gene knock-out strategies. In this study we confirm that the monocarboxylic acid transporter Jen1p is not active in 3HP. We also show that a deletion of *PRY1* increases 3HP titres by ~27% compared to the parental strain.

Chapter 5 offers conclusions made over the span of the entire thesis, highlighting the areas where more research is needed and the outlook for the future opportunities within the field.

1 Introduction

The introduction to this thesis is split into two sections. The first, Section 1.1, covers an introduction to the field of metabolic engineering, and to the studied organism of this thesis. Section 1.2 proceeds to introduce the tools and methods that have been developed within metabolic engineering and synthetic biology, and are used throughout the research section of the thesis.

1.1 Introduction to the metabolic engineering of *Saccharomyces cerevisiae*

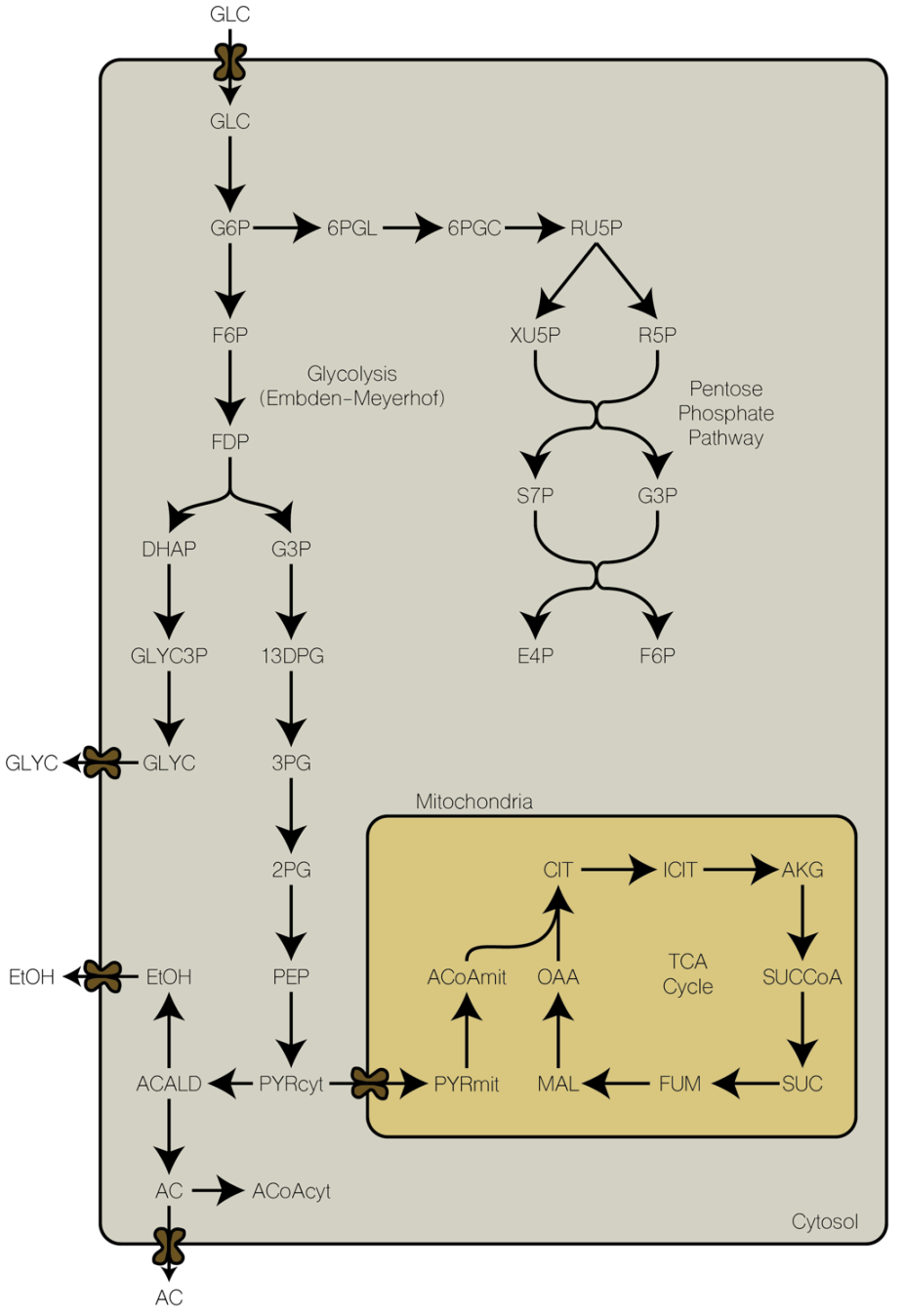
At present, the majority of chemical production stems from petroleum, with a market size of USD ~5 trillion in 2016 (https://www.statista.com/statistics/302081/revenue-of-global-chemical-industry/). The petrochemical industry contributes 18% of the global CO_2 emissions, and is reliant on a dwindling supply of finite fossil fuel reserves (Benchaita, 2013). CO_2 is the main greenhouse gas contributing to anthropogenic climate change, and as such global efforts must be made to curb such emissions (Benchaita, 2013; Solomon et al., 2009). One route that can be taken to reduce greenhouse gas emissions is through metabolic engineering. Metabolic engineering is the field of biology concerned with the alteration of the native metabolism in cells and organisms. Often within this field, single celled organisms are engineered to produce chemicals of commercial interest from carbon feedstocks such as glucose. A wide array of different chemicals can be produced through biological means, and can begin to replace chemicals that are currently produced *via* unsustainable practices. As well as the ecological benefits to bio-based production of chemicals, there is also a clear financial incentive to develop sustainable chemicals. The renewable chemicals market is predicted to show strong growth (a compound annual growth rate of 11.29%) and to be worth USD 103 billion by 2022 (https://www.zionmarketresearch.com/report/renewable-chemicals-market). The first

section of this introduction focuses on the use of the common brewer's yeast *Saccharomyces cerevisiae* as a cell factory for the renewable production of chemicals.

1.1.1 Overview of *Saccharomyces cerevisiae*

Alcoholic beverages have a long history in human societies, with evidence for fermented beverage production dating back at least 9,000 years (McGovern et al., 2004). The producers of these beverages discovered that by inoculating bread dough with small amounts of the beverage, it would cause the dough to rise. These are early examples of using microbes to produce chemicals (ethanol and carbon dioxide) that had some value to society. Yeast was later recognised as the living organism that is the source of alcoholic fermentation in beer in 1857 (Pasteur, 1995). Since then, the brewers yeast *Saccharomyces cerevisiae* (often referred to simply as "yeast") has been extensively studied (Botstein et al., 1997). It has helped shed light on many human diseases, and is routinely used in basic genetics and cancer research (Botstein and Fink, 2011; Guaragnella et al., 2014).

Figure 1 (See next page). Central carbon metabolism of S. cerevisiae. Under high concentrations, even when oxygen supply is also high, S. cerevisiae channels most of the flux from glucose into ethanol production (the Crabtree effect). Glucose catabolite repression causes the shut-down of aerobic respiration pathways and can present a problem for the implementation of pathways dependent on ATP, as not enough can be generated through the ethanol production route to provide energy for heterologous pathways. Abbreviations; GLC glucose, G6P glucose-6-phosphate, 6PGL D-6-phospho-glucono-δ-lactone, 6PGC 6-phospho-D-gluconate, RU5P ribulose-5-phosphate, XU5P xylulose-5-phosphate, R5P ribose-5-phosphate, S7P sedoheptulose-7-phosphate, G3P glyceraldehyde-3-phosphate, E4P erythrose-4-phosphate, F6P fructose-6-phosphate, FDP fructose-1,6-diphosphate, DHAP dihydroxy-acetone-phosphate, GLYC3P glycerol-3-phosphate, GLYC glycerol, 13DPG 1,3-diphosphateglycerate, 3PG 3-phosphoglycerate, 2PG 2-phosphoglycerate, PEP phosphoenolpyruvate, PYRcyt pyruvate (cytosolic), ACALD acetaldehyde, EtOH ethanol, AC acetate, PYRmit pyruvate (mitochondrial), ACoAmit acetyl-CoA (mitochondrial), CIT citrate, ICIT isocitrate, AKG α-ketoglutarate, SUCCoA succinyl-CoA, SUC succinate, FUM fumarate, MAL malate, OAA oxaloacetate.


GLC
GLC
G6P
6PGL
6PGC
RU5P
F6P
XU5P
R5P
Glycolysis
(Embden-Meyerhof)
Pentose
Phosphate
Pathway
FDP
S7P
G3P
DHAP
G3P
GLYC3P
13DPG
E4P
F6P
GLYC
GLYC
3PG
Mitochondria
2PG
CIT
ICIT
AKG
EtOH
EtOH
PEP
ACoAmit
OAA
TCA
Cycle
SUCCoA
ACALD
PYRcyt
PYRmit
MAL
FUM
SUC
AC
ACoAcyt
Cytosol
AC

S. cerevisiae is an important industrial organism due to many features that make it suitable for an array of industrial applications. Yeast holds GRAS status (Generally Regarded as Safe), allowing it to be easily and safely used in research as well as in industrial processes. In addition, yeast has been well characterised, being the first eukaryotic organism to have its genome sequenced (Goffeau et al., 1996; Nevoigt, 2008). Common laboratory strains are easy to genetically manipulate, with established techniques for inserting and deleting DNA into/from the yeast chromosome (described further in section 1.2.2). As well as chromosomal gene integration, yeast can also express genes maintained in plasmids, which makes them flexible host organisms. *S. cerevisiae* also played an important role in the creation of the first synthetic genome, as it is able to maintain and recombine large sections of heterologous DNA (Gibson et al., 2008; Kouprina and Larionov, 2016). Yeast is relatively fast growing for a eukaryotic organism, with a generation time of ~90 minutes when grown on glucose (Sherman, 2002), and several different carbon sources are able to be metabolised by yeast, such as; glucose, galactose, sucrose, fructose, maltose, ethanol, and glycerol (Gancedo, 1998; Zaman et al., 2008).

Glucose is the preferred carbon source for yeast (see Figure 1 for an overview of central carbon metabolism in yeast), and in the presence of glucose there is tight regulation of many genes and pathways involved in non-glucose metabolism (Gancedo, 1998; Trumbly, 1992). Such control usually occurs at the transcriptional level, effected largely through the Snf3p/Rgt2p glucose sensing cascade (Kaniak et al., 2004). Under high concentrations of glucose, yeast preferentially switches to fermentative metabolism even when oxygen is abundant (as discussed further in Section 1.1.2).

As a eukaryote, yeast is able to perform some functions that can not be performed by prokaryotes, such as certain post-translational modifications – which can be especially important in the production of recombinant therapeutic proteins (Mattanovich et al., 2012). Yeast fermentations can be performed under low pH conditions, which can be

advantageous in limiting microbial contamination during fermentation, and can also be beneficial in the production of organic acids where a low pH can reduce the costs associated with product separation (Baek et al., 2017; Chen et al., 2018; Curran et al., 2013). In addition, yeast is not susceptible to phage contaminations during fermentation, unlike bacterial hosts such as *Escherichia coli* (Hong and Nielsen, 2012; Runguphan and Keasling, 2014). Metabolic engineering of *S. cerevisiae* for the production of acids (in particular – hydroxy acids) is a central theme to this thesis and is covered more extensively in Section 1.1.4.

1.1.2 Crabtree-negative *S. cerevisiae*

The Crabtree effect was discovered by Herbert Grace Crabtree and is based on the observation that *S. cerevisiae* will opt to ferment glucose when it is in high concentrations and will do so even under aerobic conditions (De Deken, 1966; Pronk et al., 1996). The reasons behind this behaviour are not yet fully understood, but one explanation is that although the yield of ATP through fermentation is greatly reduced compared to respiration, the rate of ATP formation is actually higher, and so yeast may use this strategy to out-compete rivals (Pfeiffer and Morley, 2014). Another theory is that this behaviour is in fact a result of limitations in protein production and allocation, which causes metabolism to shift to pathways that are less efficient but capable of higher rates of flux that are thereby allowing the metabolism to maintain a higher growth rate (Molenaar et al., 2009; Sánchez et al., 2017). Whatever the precise reason behind this phenomenon, the practical implication is that a great amount of the carbon flux from the substrate gets converted into ethanol when fermentation is performed with a fermentable sugar such as glucose. Three strategies can be adopted to reduce or eliminate flux into ethanol;

1. Feeding with a non-fermentable carbon source such as ethanol or glycerol

2. Slow feeding of a fermentable carbon source in aerobic conditions to stay below the threshold where yeast begins to ferment
3. Engineering the yeast strain to be incapable of producing ethanol

Although the first two strategies can be of benefit to creating economical bioprocesses, they are not always suitable solutions, and therefore engineering a strain that is incapable of ethanol production has been the focus of much interest within the field. Perhaps the most obvious way to construct a Crabtree-negative *S. cerevisiae* strain is through the deletion or down-regulation of the alcohol dehydrogeneases (ADH). Drewke, Thielen and Ciriacy, (1990) successfully created a strain lacking detectable ADH activity, but high levels of ethanol were still formed, along with an accumulation of toxic acetaldehyde due to acetaldehyde being produced but unable to be converted to ethanol making this strain unsuitable as a production organism.

Another strategy, as shown in Figure 2, is to delete the three major pyruvate decarboxylase (PDC) isozymes encoded by the genes *PDC1*, *PDC5*, and *PDC6*, stopping the conversion of pyruvate to acetaldehyde (Flikweert et al., 1996, Flikweert et al., 1997). Deleting these genes leaves the strain incapable of growing on high glucose concentrations due to the effects of glucose repression, and the addition of a C_2 compound such as acetate or ethanol is required to rescue growth (Flikweert et al., 1999). A Pdc^- strain is incapable of meeting its requirements for cytosolic acetyl-CoA biosynthesis, even though the mitochondrial pyruvate dehydrogenase pathway is still functional, suggesting a lack of export mechanisms of acetyl-CoA from the mitochondria (Flikweert et al., 1999). To relieve the dependency on C_2 compounds, and to enable growth on high concentrations of glucose, directed evolution was performed on a Pdc^- strain (van Maris et al., 2004a). The strain was first gradually fed less acetate, until it could grow completely in the absence of it. It was then grown under steadily increasing concentrations of glucose, until a mutant was able to tolerate glucose concentrations of 100 g L^{-1} (van Maris et al., 2004a). Perhaps the single

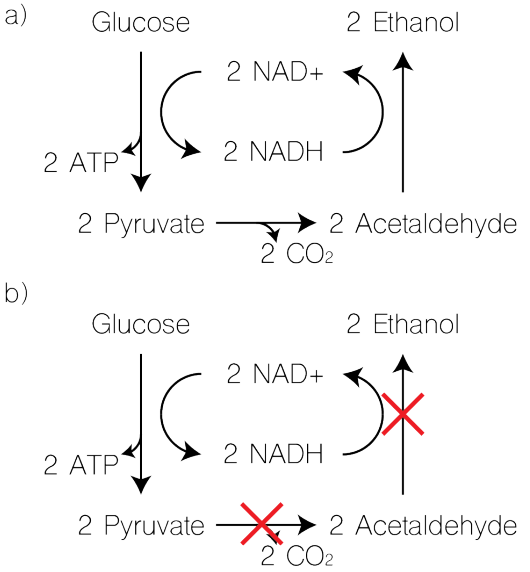


Figure 2. Ethanol production from glucose in S. cerevisiae. a) the native pathway, generating 2 mol ATP per mol of glucose and maintaining a closed redox loop by the regeneration of NAD+. b) The pathway of a S. cerevisiae strain lacking pyruvate decarboxylase activity. Ethanol is no longer produced, but neither is NAD+, leading to a redox imbalance in the cytosol.

largest contributing factor to the ability of the evolved strain (termed the TAM strain) to grow on high concentrations of glucose is a large deletion found in the glucose sensing transcriptional regulator *MTH1* (Oud et al., 2012). This deletion is hypothesised to result in a decreased degradation rate of Mth1p, which in turn restricts the import of glucose into the cell to rates at which the cell can manage which prevents the build-up of toxic levels of intracellular pyruvate (Oud et al., 2012). The restriction of glycolysis in this way is also thought to allow time for the mitochondrial respiratory chain to reoxidise NADH, resulting in a NADH/NAD+ ratio high enough to sustain growth (Oud et al., 2012). Overexpression of *MTH1* in the non-evolved Pdc$^-$ strain partially restored the ability of the strain to grow on glucose as the sole carbon source, confirming the importance of this gene to the observed phenotype of the TAM strain (Oud et al., 2012).

Cytosolic acetyl-CoA supply is also reliant on pyruvate decarboxylase activity, and a Pdc$^-$ strain is therefore unlikely to be able to produce the acetyl-CoA that is required for lipid biosynthesis – unless mitochondrial acetyl-CoA can be transported into the cytosol. The CoA transferase gene *ACH1* is thought to facilitate the indirect production of acetyl-CoA in Pdc$^-$ yeast strains. Ach1p is thought to convert acetyl-CoA to acetate, which can then be transported across the mitochondrial membrane, and then converted back to acetyl-CoA in the cytosol (Chen et al., 2015).

However, even with these compensatory pathways the TAM strain has much less robust growth characteristics than the WT strain (0.20 h^{-1} *vs.* 0.33 h^{-1}) when grown on glucose but does not produce any ethanol (van Maris et al., 2004a). Due to the lack of ethanol formation, pyruvate accumulation is observed, making it an attractive host for producing compounds derived from the central carbon metabolism such as the organic acids (van Maris et al., 2004a). Research paper 2 in this thesis explores this strain's capacity to produce various hydroxy acids derived from pyruvate, and how these modifications affect the physiology and metabolic flux distribution of the host strain.

1.1.3 Yeast as a cell factory

Aside from its role in the food and beverage industries, yeast serves as an integral part of the biotechnology industry. Routinely used in bioethanol production, it is also used in the production of other biofuels as well as; therapeutic proteins, pharmaceuticals, flavourings, fragrances, and biodegradable polymers – see Figure 3 (Jensen and Keasling, 2015; Krivoruchko et al., 2011; Li and Borodina, 2015; Nielsen, 2015). Chemical production from yeast, as with other organisms, can offer a sustainable path to replace many chemicals that are currently produced from petroleum (Borodina and Nielsen, 2014; Van Dien, 2013; Morales et al., 2016). While the environmental benefits of bio-based production are promising, the processes must still be able to compete economically with existing

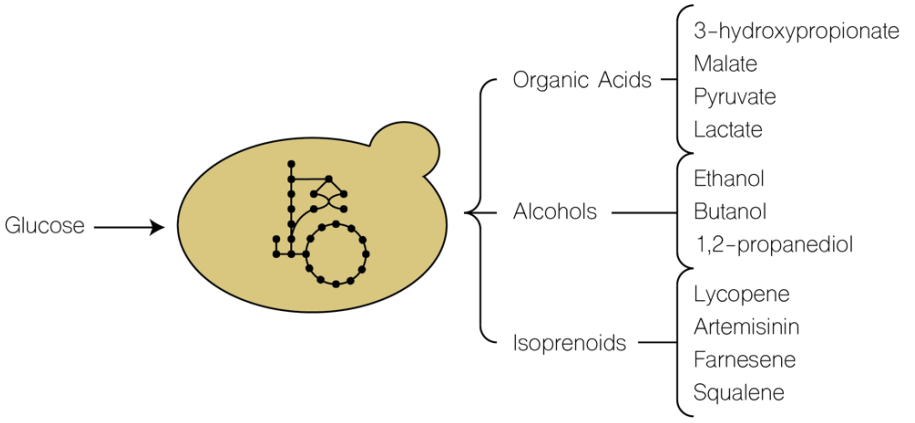


Figure 3. Yeast as a production organism. S. cerevisiae is capable of producing an array of different fine chemicals, bulk chemicals, and biofuels. Organic acids, alcohols, and isoprenoids are common classes of chemicals that are the targets of metabolic engineering efforts in S. cerevisiae.

technologies before they can be widely adopted. Commercial production of bulk chemicals from biorefineries is currently limited to a handful of cases, although projected to grow significantly in the coming years (Nielsen and Keasling, 2017; Nielsen et al., 2013). Common challenges must be addressed when producing bulk chemicals from biological hosts, such as how to divert flux from the native catabolic pathways into those that produce the target compounds.

The following section will focus on production of hydroxy acids in yeast. Hydroxy acids are bulk chemicals that need to be produced to high titres and efficiencies if they are able to compete with traditional manufacturing processes. Yeast fulfils many of the requirements to be a valuable production organism, but much research is still needed to fully utilise the potential of yeast for the efficient production of bulk chemicals (Dai et al., 2015; Fletcher et al., 2015; Nielsen et al., 2013).

1.1.4 Hydroxy acid production in yeast

Hydroxy acids are a sub-class of carboxylic acids that have a hydroxyl group located along the carbon backbone (Karlheinz, 2000). They can be further subdivided into three categories; α-hydroxy acids (AHAs) have a hydroxyl group located on the carbon atom that is one away from the carboxyl group (denoted as the α-position), β-hydroxy acids (BHAs) have a hydroxyl group on the second carbon atom away from the carboxyl group, and Ω-hydroxy acids (OHAs) have a hydroxyl group on the carbon atom that is most distant from the carboxyl group (Karlheinz, 2000; Kornhauser et al., 2010). AHAs have become well known for their use in cosmetic products, with a range of different AHAs used to chemically peel outer layers of skin (Fiume, 2017). The smallest of the AHAs is glycolic acid, one of the main AHAs used in cosmetics and healthcare, but also in many

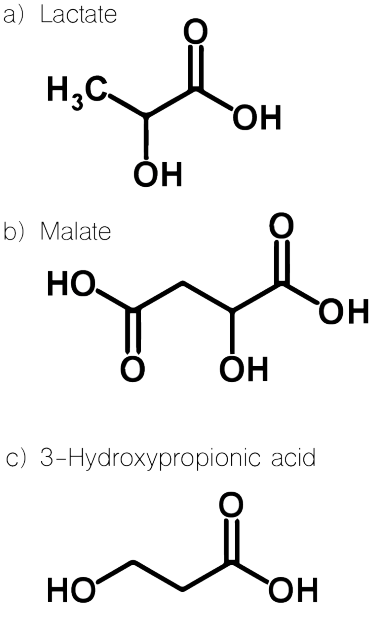


Figure 4. Chemical structures for three commonly produced hydroxyacids. Lactate (a), and malate (b) are both examples of α-hydroxy acids, with the hydroxyl group adjacent to the carboxyl group. 3-hydroxypropionic acid (c) is a β-hydroxy acid as its hydroxy group is located on the second carbon atom from the carboxyl group.

industrial processes such as in food and pharmaceutical production (Fredenberg et al., 2011; He et al., 2010; Murad et al., 1995; Reed and Gilding, 1981). Glycolic acid can be polymerised to create biodegradable plastics, either as a homo-polymer, or as a co-polymer with lactate and is therefore of interest to researchers working towards the sustainable production of plastics (Ayyoob et al., 2017; Gentile et al., 2014; Sarazin et al., 2004).

1.1.4.1 Lactate

Lactate (Figure 4a) is the next largest AHA after glycolic acid, with three carbon atoms. Like glycolic acid, lactate is widely used in the cosmetic and pharmaceutical industries. It is also used to produce polylactic acid, a biodegradable plastic used in a variety of applications from industry and healthcare, to packaging and consumer products (Chiulan et al., 2018; Jamshidian et al., 2010; Juturu and Wu, 2016; Rathin and Michael, 2006). The global market for lactate was valued at USD 1.275 billion in 2014, with a strong growth forecast (https://www.credenceresearch.com/report/lactic-acid-market). Lactate is a common metabolite found in many species as a carbon sink, and much work has been done towards optimising lactate production in a range of species (Fumi et al., 2009; Hofvendahl and Hahn–Hägerdal, 2000; Ilmén et al.; Mazumdar et al.; Meussen et al., 2012; Sauer et al., 2010). Lactate can be synthesised chemically, but in this process both of lactate's optical isomers (D and L) are produced, which is undesirable for certain applications where optically pure lactate is required (Sauer et al., 2010). Microbial fermentation can produce highly optically pure L-lactate, making it an ideal method of production (Abdel-Rahman et al., 2011; Saitoh et al., 2005; Tan et al., 2018; Yun et al., 2003). Common hosts for lactate production are the lactic acid bacteria that naturally produce high titres (>100 g L^{-1}) under fermentative conditions (Abdel-Rahman et al., 2011; Papagianni, 2012; Tan et al., 2018; Upadhyaya et al., 2014). Although capable of high titres, lactic acid bacteria based fermentations need to be performed at near to neutral pH, resulting in high separation costs, which lowers the economical efficiency of the process

(Gao et al., 2011; Sauer et al., 2010). A host able to tolerate low pH is therefore desirable for industrial lactate production, and considerable effort has been made to produce lactate in yeast (Branduardi et al., 2006; Colombié et al., 2003; Ishida et al., 2005; Sauer et al., 2010).

Lactate can be produced in a single enzymatic step from pyruvate, requiring only the reduction of the ketone functional group by a lactate dehydrogenase (LDH). Flux through pyruvate is high in yeast, as pyruvate is an intermediate in the production of ethanol (Pronk et al., 1996). Lactate production is theoretically able to replace ethanol production in yeast and act as a carbon sink, as one mole of glucose is converted into two moles of lactate and two moles of ATP, the same ratio as for ethanol production. The process also forms a closed redox loop, with the NADH formed in glycolysis being oxidised through the reduction of pyruvate, as is also the case for ethanol production. Efficient lactate fermentations have been achieved in *S. cerevisiae* strains lacking the ability to produce ethanol - see section 1.1.2 for more information on Crabtree-negative *S. cerevisiae* (van Maris et al., 2004b). A pyruvate decarboxylase-negative strain of *S. cerevisiae* engineered with the addition of an LDH gene from *Lactobacillus plantarum* has been shown to be capable of producing ~50 g L^{-1} lactate in buffered media containing 70 g L^{-1} glucose (a yield of 1.44 mol mol^{-1} glucose), highlighting the promise for the use of yeast as a host organism (Lui and Lievense, 2005). More recently, Lee et al., (2015) demonstrated a strain that was capable of producing 117 g L^{-1} lactate in a fed-batch bioreactor. This high titre was achieved through the deletion of *PDC1*, *CYB2, GPD1, NDE1,* and *NDE2* from the genome of *S. cerevisiae* with the addition of several copies of a heterologous LDH from *Pelodiscus sinensis*. Deletion of *CYB2* prevents lactate utilisation, while a *GPD1* deletion prevents glycerol synthesis. The redox state of the cell was engineered through deletion of the NADH dehydrogenases *NDE1* and *2*. The majority of NAD^+ regeneration was now a result of lactate production, enabling this strain to produce high lactate titres (Lee et al., 2015). The positive effects of *NDE1* and *2* deletions on the improvement of lactate

production in this strain are not clear, but this strategy did increase cytosolic NADH levels and so could present a target for future engineering attempts for other pathways that require redox balancing. However, this strain was unable to reach higher yields than previously achieved by Lui and Lievense, (2005), producing a yield of 0.945 mol mol^{-1} glucose in shakeflasks and 1.17 mol mol^{-1} glucose in fed-batch fermentation (Lee et al., 2015). Importantly though, these were carried out in un-buffered media which is an important step to be able to make the process economically viable as the efficient extraction of lactate from the media occurs only at low pH (Sauer et al., 2010).

To make the process more cost efficient, anaerobic fermentation is also desirable. But although homofermentative lactate production is similar in many ways to ethanol fermentation, lactate production is unable to support the anaerobic growth of yeast in the Crabtree-negative TAM strain, most likely due to ATP-dependent export of lactate (Maris et al., 2004).

1.1.4.2 Malate

Malate (Figure 4b) is a four carbon AHA, containing two carboxyl groups. It is present as an intermediate in the tricarboxylic acid (TCA) cycle, and therefore widespread across all aerobic species (Fernie et al., 2004). Malate occurs naturally in a variety of food and drinks, and is often used as an additive (Kunkee, 1968; Tsao et al., 1999). Like the previously mentioned AHAs, it is routinely used in cosmetics, pharmaceuticals, and in industrial processes, as well as in food and beverage production (Fiume, 2001; Liu et al., 2017a; Werpy and Petersen, 2004). Valued at USD 159 million in 2016, the market size is smaller than for lactate, but still represents a large economic opportunity, with growth forecast to remain strong into the next decade (https://www.grandviewresearch.com/industry-analysis/malic-acid-market).

Malate can be produced efficiently *via* a redox-neutral two-step pathway converting pyruvate first to oxaloacetate with the incorporation of CO_2 (Liu et al., 2017b; Yin et al., 2015). This reductive TCA cycle also has a maximum product yield of 2 mol mol^{-1} glucose, identical to that of the lactate pathway (Yin et al., 2015). Unlike lactate however, there is no net production of ATP, as the conversion of pyruvate to oxaloacetate requires the expenditure of one mole of ATP (Zelle et al., 2008). Therefore, even with energy neutral export processes, anaerobic production would not be possible *via* this pathway.

Aspergillus flavus is capable of naturally producing high amounts of malate, and has achieved yields of 1.26 mol mol^{-1} glucose (Battat et al., 1991). While promising as a production organism, *A. flavus* produces aflatoxins which makes it unsuitable for industrial production of malate (Do and Choi, 2007). The non-aflatoxin producing *Aspergillus oryzae* has also shown promise as a cell factory for the production of malate, with yields of up to 1.38 mol mol^{-1} glucose, and titres of 154 g L^{-1} (Brown et al., 2013). Bacterial production methods are also showing promise, with research into using *E. coli* producing high titres and yields (Guo et al., 2018; Li et al., 2018; Zhang et al., 2015). Through the rerouting of carbon metabolism to phosphoenolpyruvate *via* several gene deletions, and the use of a two stage fermentation process, malate yields of 1.42 mol mol^{-1} glucose were achieved in *E. coli* (Zhang et al., 2015). However, the titre and productivity remained low, excluding it from use in an industrial setting.

Effort has also been invested in engineering *S. cerevisiae* to produce high malate titres. Natively, *S. cerevisiae* only produces low levels of secreted malate, but through the engineering of the Crabtree-negative TAM strain (see Section 1.1.2), a yield of 0.42 mol mol^{-1} glucose and titre of 59 g L^{-1} were possible (Zelle et al., 2008). The implemented strategy involved overexpression of the pyruvate carboxylase gene *PYC2*, introduction of a variant of the malate dehydrogenase gene *MDH3* lacking the peroxisomal targeting sequence (McAlister-Henn et al., 1995), and expression of a heterologous malate

transporter from *Schizosaccharomyces pombe* (Zelle et al., 2008). To achieve these titres, $CaCO_3$ was used as a buffering agent to neutralise the excreted malate (as is also commonly used in lactate production studies). The use of a buffering agent may not be practical in industrial fermentations, and such factors should be borne in mind when comparing different production strategies. After the optimisation of the fermentation media and conditions, a malate yield of 0.48 mol mol^{-1} glucose was obtained (Zelle et al., 2010). While these titres and yields fall behind those obtained from *Aspergillus* species, it demonstrates the potential for *S. cerevisiae* to become a platform strain in malate production.

1.1.4.3 3-Hydroxypropionic acid

3-Hydroxypropionic acid (3HP) is a three carbon BHA (Figure 4c) that can be used to produce other chemicals such as acrylic acid, which alone in 2011 had a market value of USD 11.5 billion (Borodina et al., 2015). 3HP can be polymerised to form biodegradable polymers, and can be converted into acrylic acid and other compounds of industrial interest; including plastics and superabsorbent polymers (Andreessen et al., 2010; Borodina et al., 2015; Jiang et al., 2009; Li et al., 2016).

Unlike lactate and malate, 3HP has not been identified as a native highly produced compound in any organism to date, but has been shown to be a metabolic end product in glycerol fermentation by organisms such as *Lactobacillus* 208A, and *Lactobacillus reuteri* (Sobolov and Smiley, 1960; Talarico and Dobrogosz, 1990) – for a comprehensive review see: Kumar et al., (2013). It is also found in the CO_2 fixation pathway of at least one organism – *Chloroflexus aurantiacus* (Holo, 1989).

Several pathways have been evaluated as possible routes of 3HP production, and so far two main pathways have been found to be thermodynamically feasible from glucose as a substrate (Borodina et al., 2015; Kildegaard et al., 2016; Kumar et al., 2013). Of these, the

pathway with the most promising characteristics is one that goes through a β-alanine intermediate, which can have a net production of 1 mol ATP per mol of 3HP formed. Another pathway that was predicted to be thermodynamically feasible, goes through acetyl-CoA and malonyl-CoA as precursors (Figure 5). Unlike the β-alanine route, this pathway has no net production of ATP, but does have a balanced co-factor cycle. Both of these pathways (shown in Figure 5) have been successfully engineered in *S. cerevisiae,* with the β-alanine pathway proving to be capable of producing higher titres of 3HP than the malonyl-CoA pathway (Borodina et al., 2015; Kildegaard et al., 2016).

The production of 3HP *via* β-alanine (shown in Figure 5), requires the introduction of a heterologous enzyme for the conversion of aspartate to β-alanine, as the native β-alanine biosynthesis pathway is unlikely to be able to carry high levels of flux (Borodina et al., 2015). The aspartate-1-dehydrogenase gene from *Tribolium castaneum* (p*anD*) was shown to produce β-alanine from aspartate, with a β-alanine pyruvate aminotransferase (BAPAT) from *Bacillus cereus* capable of the inter-conversion of β-alanine and pyruvate to L-alanine and malonic semialdehyde. The final step from malonic semialdehyde to 3HP is catalysed by a NADPH dependent 3HP dehydrogenase (HPDH) from *E. coli* (*ydfG*). Construction of this pathway in yeast resulted in a strain that produced ~14 g L^{-1} 3HP in a fed-batch fermentation, representing a product yield of 0.27 mol mol^{-1} glucose (Borodina et al., 2015).

Through the recent construction of a kinetic model of the β-alanine route of 3HP biosynthesis, it is predicted that 3HP titres can be increased through the overexpression of the final two enzymes in the pathway (BAPAT and HPDH) (Dalwadi et al., 2017). Overexpression of just one of these would not likely lead to an increase in 3HP titres, and if only BAPAT was overexpressed there would be an accumulation of malonic semialdehyde, a toxic intermediate.

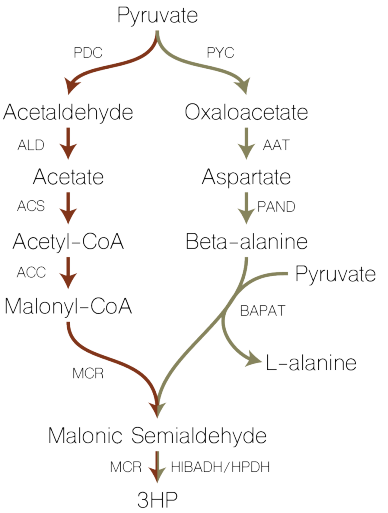


Figure 5. Two routes to 3HP production in yeast. The red pathway summarises the route via malonyl-CoA, and the green pathway is the route via β-alanine. The β-alanine route has been shown to be more efficient for 3HP production in yeast, but optimisation is still necessary to achieve commercial titres, rates, and yields. For each step the enzyme responsible for the catalysis of the reaction is shown. Metabolite abbreviations; 3HP – 3-hydroxypropionic acid. Enzyme abbreviations; PDC – pyruvate decarboxylase, ALD – acetaldehyde dehydrogenase, ACS – acetyl-CoA synthase, ACC – acetyl-CoA carboxylase, MCR – malonyl-CoA reductase, PYC – pyruvate carboxylase, AAT – aspartate aminotransferase, PAND – aspartate decarboxylase, BAPAT - β-alanine pyruvate aminotransferase, HIBADH – 3-hydroxyisobutarate dehydrogenase, HPDH – 3-hydroxypropionate dehydrogenase.

The production of 3HP through malonyl-CoA, while capable of lower theoretical maximum yields than the β-alanine route, can still present a viable pathway for high titre 3HP synthesis (Kildegaard et al., 2016). *C. aurantiacus* synthesises 3HP from malonyl-CoA, requiring 2 moles of NADPH per mole of malonyl-CoA *via* malonyl-CoA reductase (MCR), and this enzyme can be functionally expressed in yeast (Kildegaard et al., 2016; Tabita, 2009). Due to the net zero ATP production of this pathway, high 3HP yields are more heavily dependent on the TCA cycle and efficient aeration than in the β-alanine

pathway, which could limit 3HP titres when scaling up fermentations (Borodina et al., 2015). This pathway is dependent on a plentiful supply of cytosolic acetyl-CoA, and optimisation of acetyl-CoA supply strategies has been shown to have a strong impact on 3HP titres (Jessop-Fabre et al., 2016 - Research article 1 in this thesis). Acetyl-CoA is also important in the production of a range of other chemicals, and so it is necessary to engineer robust supply pathways that serve a broad range of production pathways (Nielsen, 2014).

Engineering of the the pyruvate dehydrogenase (PDH) bypass can be a successful strategy to increase pools of acetyl-CoA (Shiba et al., 2007). Overexpression of the native pyruvate decarboxylase gene *PDC1*, and aldehyde dehydrogenase gene *ALD6*, along with expression of a heterologous acetyl-CoA synthase (ACS) proved a robust strategy in the production of isoprenoids dependent on acetyl-CoA (Shiba et al., 2007). Native acetyl-CoA synthase (Acs1p) is negatively repressed by glucose, keeping cytosolic acetyl-CoA pools low (Nielsen, 2014). The use of a *Salmonella enterica* ACS that is not sensitive to the feedback inhibition mechanisms of yeast can increase flux into acetyl-CoA production (Shiba et al., 2007). Conversion of acetyl-CoA to malonyl-CoA is natively catalysed by the acetyl-CoA carboxylase Acc1p (Lynen, 1959; Wakil et al., 1983). This step is also subject to glucose repression, with Acc1p being controlled at both the transcriptional and post-translational levels (Shi et al., 2014; Shirra et al., 2001; Tehlivets et al., 2007). Post-translational control of Acc1p is effected through the glucose repressor Snf1p. However, deletion of *SNF1* is not a viable approach to lift glucose repression of Acc1p, as Snf1p plays a role in various processes that are controlled by glucose responsive pathways (Hedbacker and Carlson, 2008; Thompson-Jaeger et al., 1991; Wilson et al., 1996). Instead, site directed mutagenesis (SDM) has been used to produce mutant versions of Acc1p that are no longer subject to Snf1p repression (Shi et al., 2014). A double mutant created by SDM , $ACC1^{S659A, S1157A}$ is particularly beneficial to the production of fatty acid ethyl esters and 3HP, and has been used successfully in further metabolic engineering studies (Jessop-

Fabre et al., 2016 - Research article 1 in this thesis; Kildegaard et al., 2016; Li et al., 2015; Shi et al., 2014).

Introduction of a cytosolically localised PDH complex can circumvent the increased production of acetaldehyde (a toxic intermediate) associated with the PDH bypass by producing acetyl-CoA directly from pyruvate. In yeast the PDH complex is capable of performing this enzymatic step, but is located only in the mitochondria. However, the successful expression of a PDH complex from *Enterococcus faecalis* into the cytosol of *S. cerevisiae* has been previously demonstrated (Kozak et al., 2014). This complex is large, with an estimated mass of 14 MDa (Snoep et al., 1992). Unlike production through ACS, this enzyme does not require ATP to convert pyruvate into acetyl-CoA and is also insensitive to high NADH/NAD^+ ratios (Kozak et al., 2014). However, the correct functionality of this enzyme is dependent on lipoic acid supplementation. This enzyme effectively boosted final 3HP titres by nearly 100% in a *S. cerevisiae* strain expressing the malonyl-CoA pathway (Jessop-Fabre et al., 2016 - Research article 1 in this thesis).

1.2 Tools for the metabolic engineering of yeast

Progress towards the goals of a metabolic engineering project is dependent on the tools available for designing, building, testing, and learning from new strategies. This "Design-Build-Test-Learn" cycle (as illustrated in Figure 6) is central to all engineering practices. The second half of the introduction chapter will focus on the tools and methods used throughout the research sections of this thesis.

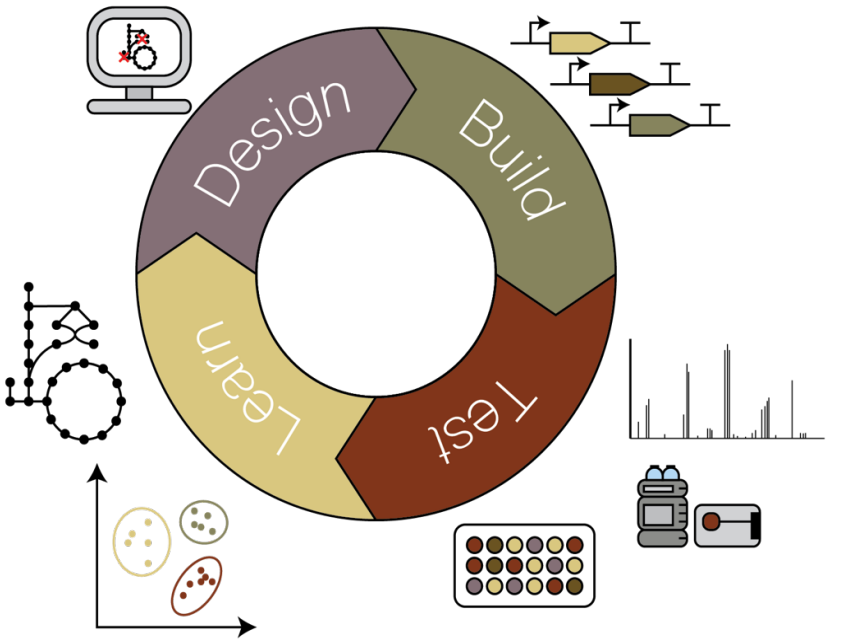


Figure 6. The "Design-Build-Test-Learn" cycle adopted in metabolic engineering. The cycle is fundamental to metabolic engineering practices. Strain design is becoming more and more automated by the development of software capable of predicting successful production strategies. The build part is becoming increasingly high-throughput, with technologies being developed to create entire libraries of strains rather than producing them one by one. Strain testing is largely reliant on the use of chromatographic and mass spectrometry methods, which are incapable of dealing with the large number of strains that can be developed.

1.2.1 Design

1.2.1.1 Metabolic modelling

Computer aided design (CAD) is commonly used in many fields of engineering, and so too is it used in metabolic engineering (Cardoso et al., 2018; Delepine et al., 2016; Fehér et al., 2014; Machado and Herrgård, 2015). For CAD to be a useful tool to the biological engineer, there need to be sophisticated models of the chosen organisms and techniques for analysing these models and to generate predictions for successful strain design strategies. One way in which it is possible to create a model of an organism is by viewing cellular metabolism as a set of equations. In the most basic form, each enzymatic reaction can be expressed in terms of a mass balance *e.g.* for the reaction A + B → C, there will be a net consumption of one mole of A, and one mole of B, with a net production of one mole of C in each cycle. In the reaction; 2D → B + E, there will be a net consumption of two moles of D with a net production of one mole of B and one mole of E. Each reaction in a cell can be formalised in this manner, and can be built up into a stoichiometric matrix

$$\begin{array}{c} \\ m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_n \end{array} \begin{array}{c} \begin{array}{cccccccc} r_1 & r_2 & r_3 & r_4 & . & . & . & r_n \end{array} \\ \begin{pmatrix} 1 & 0 & -1 & 0 & . & . & . & 0 \\ 2 & 0 & -1 & 0 & . & . & . & 0 \\ 0 & 2 & 0 & 1 & . & . & . & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & & \vdots \\ 0 & -1 & 0 & 1 & . & . & . & 0 \end{pmatrix} \end{array} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix}$$

*Figure 7. The stoichiometric matrix. The stoichiometry of each reaction in a system. A cellular network can be captured in a stoichiometric matrix (S-matrix). The **r** columns, represent each reaction in the system, whereas the **m** rows represent each metabolite in the system. Each reactions rate of flux under steady-state conditions is captured in the **v** variables.*

that captures this information. An example of a stoichiometric matrix is shown in Figure 7, where **r** represents a reaction, **m** represents a metabolite, and **v** represents the flux through a given reaction under steady-state conditions. Stoichiometric matrices can be used to calculate rates for each reaction in the matrix. However, several constraints have to be first set in order for the system to return feasible solutions. Thermodynamic constrains, as well as experimentally derived insights into metabolic pathways can help constrain the solution space into one that is feasible and computationally efficient (Orth et al., 2010). Figure 8 depicts a system of three different reactions (the x, y, and z axes). The unbounded model is shown in Figure 8a, where any solution is feasible, the solution space is infinite, and therefore unable to be computed. Figure 8b illustrates how each reaction can be constrained, limiting the available solutions to being within the 'flux cone'. Furthermore, the optimal solutions exist on each vertex of the flux cone and these optimal solutions can be efficiently found using developed techniques such as flux balance analysis (FBA), which uses mixed integer linear programming to optimally solve the matrix according to an objective function, such as maximisation of biomass production (Baart and Martens, 2012; Orth et al., 2010). Parsimonious flux balance analysis (pFBA) builds upon regular FBA by finding solutions that have the lowest net fluxes, resulting in the most efficient pathways being found, and results closer to experimentally observed data are returned (Lewis et al., 2010).

Metabolic models of cellular metabolism are used to model cells *in silico* and can be used to simulate how genetic perturbations may affect the phenotypic landscape of a host strain. Once experimental data, such as protein or flux data, have been obtained, these can be applied to the model, further improving the accuracy of that model for subsequent predictions (Kim et al., 2017; Orth et al., 2010).

Genome-scale metabolic models (GEMs) can vary heavily in both size and quality, potentially limiting their applicability in metabolic engineering (Kim et al., 2017). The first

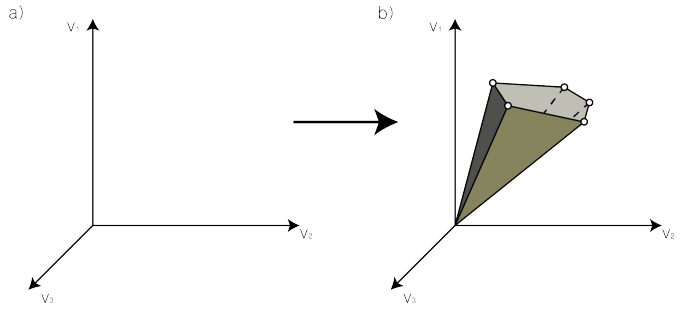


Figure 8. Constraining the solution space. An unconstrained S-matrix (a) will have an unlimited number of feasible solutions. By constraining the rates of fluxes for some reactions, the feasible solution space becomes solvable (b). Optimal solutions exist on each vertex of the 'flux cone', with linear programming algorithms capable of finding the solutions that optimise the objective function (e.g. biomass production). Adapted from Orth et al. (2010).

GEM to be constructed for *S. cerevisiae* covered 1,175 metabolic reactions, and 708 open reading frames (Förster et al., 2003). After a decade of improvements, the Yeast7 model increased the total metabolic reaction count to 3,493 reactions (Aung et al., 2013). However, an increase in captured metabolic reactions does not necessarily mean a higher quality model. One issue faced by GEMs is that they do not take into account the protein allocation or enzymatic constraints that cells are faced with. Progress towards integrating such constraints has led to models capable of more accurate predictions of cellular growth and phenotype (Lerman et al., 2012; Sánchez et al., 2017). A recently built enzyme constrained version of the Yeast7 model was able to achieve predictions that were much closer to experimental results than the unconstrained model (Sánchez et al., 2017). Using a standard GEM, it is not possible to capture the Crabtree effect in yeast (for more detail on the Crabtree effect see section 1.1.2). Incorporating enzymatic constraints into a GEM renders it able to resolve this phenomenon and accurately predict the conditions necessary for the switch from aerobic respiration to fermentation (Sánchez et al., 2017).

With a functioning GEM, computational tools can be used to predict strategies for the rational design or alteration of production strains. Many tools have been developed for the computational design and prediction of heterologous pathways, and while out of scope for this thesis, a detailed review of these tools can be found in (Wang et al., 2017). Tools also exist for the prediction of gene deletion strategies. One commonly used tool is OptKnock, that predicts gene knock-outs that can be beneficial to optimising a given objective function *e.g.* production of a compound of interest (Burgard et al., 2003). Soon after the introduction of this method, Fong et al., (2005) developed lactate producing *E. coli* strains based on OptKnock predictions, which after adaptive evolution were capable of producing up to 1.75 g L^{-1} lactate from 2 g L^{-1} glucose. In a more recent study, OptKnock was used to construct a strain of *E. coli,* where the production of the compound of interest (1,4-butanediol) was linked to the growth of the host organism. This successful strategy resulted in a strain producing 18 g L^{-1} of 1,4-butanediol (Yim et al., 2011). Such a tool can hopefully help to minimise the number of designs that are constructed in the wet lab, and therefore help to speed-up the "Design-Build-Test-Learn" cycle. Research article 3 of this thesis uses the OptKnock algorithm to design gene deletion strategies in a yeast strain capable of high 3HP titres.

1.2.1.2 Biological "Parts"

The following is one loose definition of synthetic biology; "Synthetic biology is the engineering of biology: the synthesis of complex, biologically based (or inspired) systems, which display functions that do not exist in nature", (NEST, 2005). It is seen as a field that has engineering principles at its core, and attempts to apply these engineering principles in practice (Endy, 2005). Analogies may be made to the field of electrical engineering. Both fields deal with circuits, whose elements augment the behaviour of said circuit. Electrical engineers rely on having characterised parts that can be used interchangeably and whose performance in a new environment is predictable. These parts, and the relationships

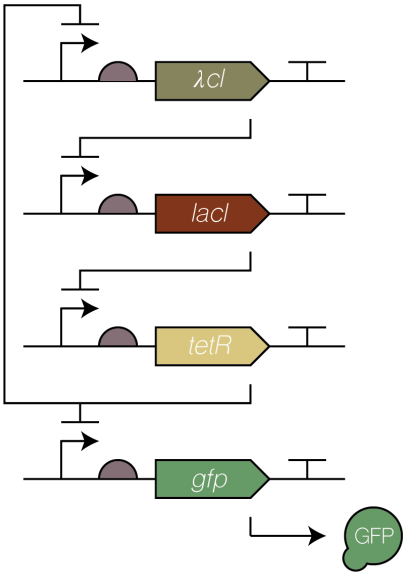


Figure 9. An example of a common "gene circuit". The repressilator circuit introduced by Elowitz and Liebler (2000). Three genetic elements in turn repress each other, with tetR also repressing the output signal (GFP). This circuit was modelled in silico, and tested in vivo, where it showed oscillatory behaviour. The diagram is drawn using the standards of the synthetic biology open language. Genetic 'parts' are represented by different symbols, and so a genetic circuit can be easily understood by anyone familiar with the SBOL convention. In this simple scheme, the bent arrow denotes the promoter region, followed by the semi-circle of the ribosomal binding site. The thick coloured arrows represent coding sequences (CDS), and the T shape at the end denotes a terminator sequence. Lines from the CDS to promoters represent the repression of the promoters by the gene products. The output macromolecule (GFP) is denoted by the circular shape.

between them can be represented visually and in a standardised format, so that an engineer on one side of the world can quickly understand and follow what an engineer has constructed on the other side of the world. In a similar way, synthetic biology aims to create standardised biological "parts" and introduce standardised methods to visualise biological circuits. Such standardisation is important in metabolic engineering, as it can help increase the reproducibility within the field, a common issue across the biological

sciences (Baker, 2016; Begley and Ellis, 2012; França and Monserrat, 2018; Sene et al., 2017). Figure 9 shows the design of a repressilator circuit (Elowitz and Leibler, 2000), one of the earliest examples of synthetic biology, as an example of how biological circuits can be visualised according to the synthetic biology open language (SBOL) (Quinn et al., 2015). Such a genetic circuit can be read an understood, and then implemented by a biological engineer in a different lab. Attention has also been paid towards improving the standardisation of the physical biological parts, often referred to as biobricks. The Registry of Standard Biological Parts attempts to catalogue and define biobricks, with each preferably having a detailed specification sheet outlining their characteristics (http://parts.igem.org/Main_Page). Basic biobricks for the engineering of yeast are promoter, gene, and terminator sequences, that can then be used interchangeably and have relatively well defined characteristics.

Some applications of synthetic biology for metabolic engineering include; genetic manipulation of host genomes, the creation of synthetic promoter libraries for tuneable pathway control, biosensor development to aid in the rapid screening of production strains, and dynamic control of production pathways (Curran et al., 2014; DiCarlo et al., 2013; Maury et al., 2018; Skjoedt et al., 2016). Synthetic biology tools for the manipulation of the yeast genome are discussed further in the following section.

1.2.2 Build

1.2.2.1 Genetic manipulation of yeast

Natural *S. cerevisiae* isolates display a wide variation in ploidy, with the majority being diploid but with several isolates having been reported to be haploid, triploid, tetraploid, or higher (Codon et al., 1998; Peter et al., 2018). They can also exhibit degrees of aneuploidy, leading to complications in their genetic manipulation. Laboratory strains used in

metabolic engineering are therefore commonly haploid and, unless stated otherwise, the following techniques are most commonly applied to haploid laboratory strains of yeast. A common method for the expression of genes in yeast is to clone the genes into a plasmid, which can then be maintained and expressed in *S. cerevisiae* (Gunge, 1983; Polumienko et al., 1986; Stinchcomb et al., 1982). Plasmids can have different origins of replication which determine whether the plasmid is kept in high or low copy numbers within the cell. For example, plasmids containing the 2μ origin are maintained at high copy numbers, while plasmids with the CEN/ARS origin of replication are maintained at lower copy numbers (Chan et al., 2013; Karim et al., 2013). Plasmid based expression can be noisy, with heterogeneous populations producing different levels of protein expression based on variation in plasmid copy numbers between individuals (Jensen et al., 2014; Karim et al., 2013). Integration of genetic material into the chromosomes of yeast at defined loci can produce populations that are far more homogenous than plasmid based expression (Jensen et al., 2014). With this technique, the copy number of each inserted gene is pre-defined making it simpler to compare results across strains – apart from when insertion into a yeast retrotransposon is performed to simultaneously introduce genes in multiple copies into the genome (Jensen et al., 2014; Maury et al., 2016).

Several tools for manipulation of the chromosomal genetic material of *S. cerevisiae* have been well established (Gietz and Schiestl, 2007; Guldener et al., 1996; Jensen et al., 2014; Da Silva and Srikrishnan, 2012). These tools rely on introducing DNA containing a selectable marker into the yeast genome *via* its native homologous recombination machinery. By having upstream and downstream regions that are homologous to the site of insertion, it is possible to precisely target the locus where the DNA should be incorporated. The inclusion of loxP sites into these sequences confers the ability of marker cassettes to be looped out of the genome, enabling the markers to be used once more in further transformations (Sauer, 1987). Two different types of marker cassettes may be used in yeast, either dominant or auxotrophic markers. Dominant markers confer antibiotic

resistance to the host strain, such as the commonly used kanMX cassette whose gene product protects cells against the action of the gentamicin analogue G418 (Wach et al., 1994). This has been further developed to contain loxP sites to enable Cre recombinase mediated excision (Guldener et al., 1996). This dominant marker cassette is routinely used for generation of gene deletion mutants, and was used to generate the yeast knock-out collection (Giaever et al., 2002).

Auxotrophic markers, on the other hand, aim to confer the ability to synthesise amino acids that the parental strain is otherwise unable to synthesise. Selection is achieved by growing the transformed strain on plates lacking the relevant amino acids. Histidine autotrophy can be selected for by using the *HIS3* gene from *Schizosaccharomyces pombe* (Wach et al., 1997). Uracil and leucine autotrophies can be selected for by insertion of *URA3* and *LEU2* from *Kluyveromyces lactis* (Gueldener et al., 2002). Auxotrophic markers allow for multiple markers to be used without requiring the markers to be looped out between transformations. However, the use of auxotrophic markers can cause issues that may impact on the experimental results. Firstly, the strain must have been engineered to be auxotrophic for specific amino acids. This is trivial when working with common lab strains where auxotrophy has already been established, but becomes a severe drawback when working with new, prototrophic, or non-haploid strains (Vickers et al., 2013). There can be further problems when screening for phenotypes, as the insertion of these selection genes can alter host physiology, potentially leading to unreliable results (Alam et al., 2016; Baganz et al., 1998; Çakar et al., 1999; Gonzalez et al., 2008; Leng and Song, 2016).

Recently, the development of clustered regularly interspaced short palindromic repeats (CRISPR) technologies have enabled yeast strains to be constructed without inserting markers into the genome, circumventing some of the restraints of traditional methods for genomic integration. A double stranded break (DSB) introduced by the CRISPR-associated protein 9 (Cas9) acts as a negative selection system since *S. cerevisiae* cannot

survive a DSB unless it is repaired (Jakočiūnas et al., 2015; Jessop-Fabre et al., 2016 - Research article 1 in this thesis; Mans et al., 2015; Reider Apel et al., 2017; Stovicek et al., 2015, Stovicek et al., 2017). This cheap and efficient technology allows for precise cutting at user defined sequences, and has seen widespread adoption within synthetic biology and metabolic engineering (Cho et al., 2018; Lian et al., 2018).

CRISPR-Cas systems were discovered in *E. coli* and have since been shown to be a bacterial and archaean adaptive defence system protecting them against viral infections (Garneau et al., 2010; Ishino et al., 1987). The CRISPR repeat sequences are separated by spacer regions often containing DNA sequences of phage origin, giving rise to the theory that it may be related to immunity of some kind (Bolotin et al., 2005; Mojica et al., 2005). In brief, the system functions by having a DNA sequence, *e.g.* from a phage, inserted into

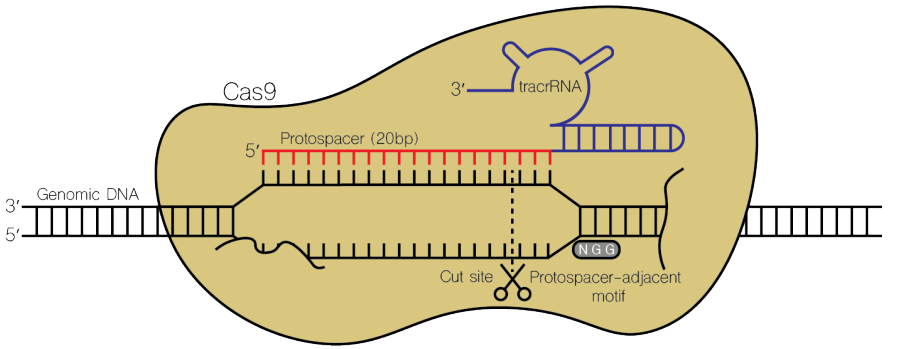


Figure 10. Mechanism of CRISPR-Cas9. The Cas9/sgRNA complex binds to the 20bp recognition sequence that is complementary to the sequence carried on the sgRNA. The presence of a PAM site upstream of the DNA/RNA binding region then promotes the Cas9 nuclease to introduce a double stranded break in the DNA. In S. cerevisiae, a double stranded break in the chromosome is lethal and so must be fixed. The presence of a length of DNA with homology regions to either side of the break can be used by the native homologous recombination machinery to repair the DSB. Abbreviations; Cas9 – CRISPR associated protein 9, tracrRNA – trans-activating CRISPR RNA.

the CRISPR locus during an infection. Afterwards, when a cell line that has previously encountered the phage gets infected again by the same phage, the DNA from the CRISPR loci are transcribed into RNA molecules, which are then matured and associate with the Cas9 protein. If the RNA molecule matches a sequence on the phage DNA, the Cas9 complex can then bind and introduce a DSB into the phage DNA, preventing the cell from further infection (Marraffini, 2015; Rath et al., 2015).

The RNA that binds to the Cas9 protein consists of a structural region, recognised by the Cas9 protein, and a recognition sequence that binds the foreign DNA (Figure 10). In the type II CRISPR system, this RNA is expressed in two parts; the CRISPR RNA (crRNA) that contains the DNA recognition sequence, and the transactivating CRISPR RNA (tracrRNA) necessary for correct binding to Cas9 (Deltcheva et al., 2011; Jinek et al., 2012). To make this system into one that is simple to port to different host organisms for synthetic biology purposes, the crRNA and tracrRNA regions from *Streptococcus pyogenes* were fused to create a single guide RNA (sgRNA) (Jinek et al., 2012).

Important in the CRISPR-Cas9 system is the requirement for the protospacer adjacent motif (PAM). This is a three base long sequence located immediately after the recognition sequence and for the *S. pyogenes* system this sequence is NGG. Without it, Cas9 will not introduce a DSB at the recognition site. CRISPR-Cas9 has become an established tool in the genetic manipulation of *S. cerevisiae* with methods developed early for the functional expression of gRNAs and Cas9 (DiCarlo et al., 2013). This method has allowed for quick, cheap, and precise creation of DSBs along the yeast genome and many tools have been developed that allow for the introduction, or removal of genes to/from the yeast genome (Jakočiūnas et al., 2015; Jessop-Fabre et al., 2016 - Research article 1 in this thesis; Mans et al., 2015; Reider Apel et al., 2017; Vanegas et al., 2017). The recognition sequence of the sgRNA (20bp) is the only sequence that needs to be modified depending on the target of interest and so can be introduced easily through common cloning procedures. The relative

ease, low-cost, and speed of creation of vectors targeting new sites are some of the biggest factors contributing to the popularity of the technology, with CRISPR-Cas9 quickly having overtaken the two other programmable nuclease technologies – transcription activator-like effector nucleases (TALENs) and zinc finger nucleases (ZFNs) (Figure 11). This thesis contributes to the development of CRISPR-Cas9 technologies is yeast, through the creation of a vector set for CRISPR-Cas9 mediated genomic integrations in research article 1.

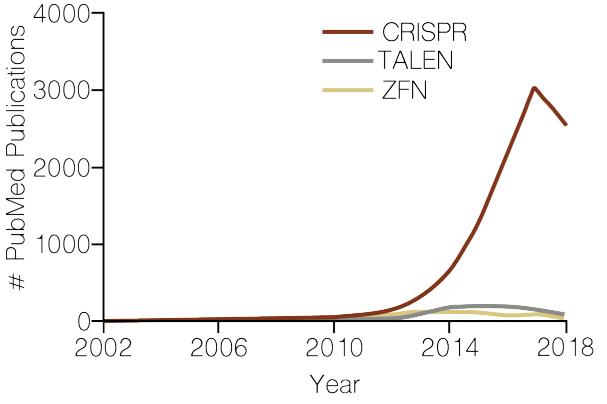


Figure 11. The popularity of CRISPR based DNA editing. CRISPR based technologies have been widely adopted, and have quickly become more popular than the two other most popular programmable nuclease technologies; Transcription activator-like effector nucleases (TALEN) and zinc finger nucleases (ZFN). Figures are generated from a PubMed search for articles containing the keyword in their titles or abstracts. 2018 totals are for the year to August.

1.2.3 Test/Learn – Omic technologies

Systems biology aims to understand biological phenomena from a top-down approach, and as such is reliant on experimental techniques capable of generating organism-level data sets. Such tools now exist and can be collected under the "-omic" technologies that cover

the range of different hierarchies present in a cell. From the genome, to transcriptome, to proteome, to fluxome, and metabolome (and many more), each can be measured quantitatively to provide an understanding of cellular phenomena. Two of these technologies relevant to this thesis, transcriptomics and fluxomics, are introduced in further detail in the following sections.

1.2.3.1 Transcriptomics

Transcriptomic technologies are used to measure the gene transcript levels in a cell or population at a specific point in time. By isolating RNA, and reverse transcribing it into complementary DNA (cDNA) it can be read *via* microarray or next-generation sequencing (NGS) platforms, with transcriptional abundances calculated (Ozsolak and Milos, 2010). Transcriptomic analysis has the ability to shed light on a strain's regulatory network, and how this network reacts once perturbed by the insertion/deletion of genes, by specific growth conditions, or by the presence or absence of specific metabolites. Such information can be used to identify where there might be opportunities for further strain improvement. Transcriptomic technologies are widely used within metabolic engineering and have contributed important knowledge to the field. However, one limitation of transcriptomics is the lack of correlation between transcript levels and the levels of its associated protein *in vivo* (Bai et al., 2015). In this way, transcriptomics becomes limited to providing information on transcriptional abundances and regulation, whereas other methods, such as proteomics or fluxomics, must be employed to quantify changes in protein abundances or fluxes, respectively.

Transcriptomic analysis is routinely used in metabolic engineering, often helping to characterise the stress responses of production organisms. In a recent study into the transcriptome of a mutant strain of *Propionibacterium acidipropionici*, Guan et al., (2018) reported that this strain is able to withstand higher than wild-type (WT) levels of propionic acid, and transcriptomic profiling resulted in the discovery of the potential mechanisms

for acid tolerance. An array of different enzymes and transcriptional regulators related to lactate and acetate production have strongly downregulated expression profiles compared to the WT strain. Deletion of the pyruvate oxidase gene *poxB* in the WT strain significantly decreased the levels of acetate formed, and increased propionic acid titres by 19.1% (Guan et al., 2018).

Transcriptomics has also been used to study the stress responses of *E. coli* in response to high levels of 3HP. Under 3HP stress *E. coli* adopts the arginine tolerance mechanism, which helps to boost the intracellular pH levels to try and combat the low extracellular pH caused by 3HP (Yung et al., 2016). Such knowledge may help researchers develop production strains showing higher tolerances to 3HP and other acids.

A combination of genomic and transcriptomic techniques can be powerful when analysing mutant strains. Adaptive laboratory evolution has generated yeast mutants that are capable of growing in up to 50 g L^{-1} 3HP (Kildegaard et al., 2014). Through transcriptomic analysis, many oxidative stress genes were found to be highly upregulated in the mutant strains, suggesting the importance of these pathways in 3HP detoxification. Two genes were found to be mutated in all three evolved strains, with non-silent point mutations in genes *ACH1* and *SFA1*. The mutated *SFA1* appears to confer a majority of the observed 3HP tolerance, with the replacement of the WT *SFA1* with $SFA1^{C276S}$ or $SFA1^{M283I}$ enabling the strains to grow in 40 g L^{-1} 3HP (Kildegaard et al., 2014). 3HP detoxification in the evolved strains is thought to occur through a similar process to that of formaldehyde detoxification, a pathway that Sfa1p is known to participate in (Kildegaard et al., 2014; Mack et al., 1988; Wehner et al., 1993).

The impact of the malonyl-CoA 3HP production pathway (Figure 5) on the transcriptome of *S. cerevisiae* revealed that the ATP and redox cofactor requiring pathway caused large transcriptional changes in the engineered strain compared to the non-producing WT strain

(Kildegaard et al., 2016). The pentose phosphate (PP) pathway, and the TCA cycle were both upregulated in the engineered strain suggesting compensation for the increased ATP and NADPH requirements.

1.2.3.2 ^{13}C-based metabolic flux analysis (Fluxomics)

^{13}C-based metabolic flux analysis (MFA) is the method of quantifying fluxes through metabolic pathways *via* the uptake of an isotopically labelled substrate. A commonly used substrate for this technique is isotopically labelled glucose. Glucose contains six carbon atoms, and can be synthesised so that individual, a combination, or all of these carbon atoms are the ^{13}C isotope instead of the natural ^{12}C. Labelled glucose is then fed to yeast during cultivation, which assimilates and converts it into metabolic intermediates in the same manner as it would naturally labelled glucose. These metabolic intermediates can themselves be measured by mass-spectrometry or nuclear magnetic resonance (NMR), with the resulting distribution of isotopomers (isotopic isomers) being used to constrain a stoichiometric model of the central metabolic pathways (Stephanopoulos, 1999; Yang, 2013; Zamboni et al., 2009). Proteinogenic amino acids can also be used to determine the central carbon metabolism fluxes since the carbon building blocks for proteins are derived from central carbon metabolism. The amino acid isotopomer distribution can be measured through gas chromatography-mass spectrometry (GC-MS). The resulting isotopomer ratios are then combined with a stoichiometric matrix of the reactions of central carbon metabolism to accurately predict through which pathways the amino acid precursors were generated from, and therefore the flux distribution (Antoniewicz, 2015; Nanchen et al., 2007; Sauer, 2004; Yang, 2013; Zamboni et al., 2009). With this approach, it is possible for metabolic engineers to analyse how fluxes are distributed through their pathways, where flux control points may occur, where toxic intermediates may be accumulating, or where competing reactions are drawing flux away from the engineered pathways (Kim et al., 2008; Lee et al., 2011).

^{13}C- MFA has provided insights into hydroxy acid production in yeast, such as a study into the malonyl-CoA route for 3HP production by (Kildegaard et al., 2016). Engineering of this pathway enabled a strain to produce up to 13 g L^{-1} 3HP with a C-mol yield of 0.26 mol mol^{-1} glucose when cultivated in a bioreactor. Through combining ^{13}C-MFA with genome scale modelling, flux distributions were analysed in more depth than is possible using only MFA. The TCA cycle of this engineered strain had an increased relative flux compared to the WT of 12%, and an increased flux of 23% through the PP pathway, with both of these pathways regenerating NADPH that can be used for 3HP synthesis. However, even with this increased activity and the overexpression of a heterologous $NADP^+$-dependent glyceraldehyde-3-phosphate the attempts to balance the redox state of the producing strain were not fully successful. Although the NADPH/$NADP^+$ ratio increased, so did the NADH/NAD^+ ratio, with glycerol production being observed, opening up new directions and targets for future redox engineering efforts (Kildegaard et al., 2016).

In an MFA study on lactate production, engineered Crabtree-negative (*PDC1* and *PDC5* deletion) and Crabtree-positive lactate producing yeast strains have been analysed under micro-aerobic conditions (Nagamori et al., 2013). Under these conditions, both the Crabtree-negative and Crabtree-positive strains were estimated to maintain similar flux distributions across upper central carbon metabolism, but begin to diverge at the pyruvate node, where the Crabtree-negative strains are able to divert more flux into lactate rather than acetaldehyde production (Nagamori et al., 2013). Under these micro-aerobic conditions, none of the tested strains showed appreciable flux through the TCA cycle, nor through the PP pathway.

Zelle et al., (2008) performed ^{13}C-MFA on a malate producing Crabtree-negative strain, derived from the TAM strain (see section 1.1.2 for more information on the TAM strain). The details of this strain design and performance can be found in section 1.1.4.2. The

malate producing strain was verified to be only producing relevant levels of malate through the engineered non-oxidative pathway, and not through the oxidative TCA cycle. The non-oxidative pathway in this strain contributes the bulk of malate production, while a significant proportion of the carbon flux is still lost through pyruvate secretion, and cytosolic fumarate production that is transported into the mitochondria. However, the estimated high conversion of malate into fumarate in the yeast cytosol may not be an accurate reflection of the actual flux distribution, as cytosolic fumarase is believed to only convert fumarate to malate (Pines et al., 1996).

The following research articles aim to contribute to the topics outlined in this introduction. Research article 1 has a focus on developing CRISPR-Cas9 tools for yeast. This work presents a toolkit for the insertion of DNA into the yeast chromosome at the specific loci of the EasyClone system. This paper also investigates some mechanisms of cytosolic acetyl-CoA production and the impact of these strategies on the production of 3HP *via* the malonyl-CoA route. The expression of a bacterial PDH complex from *E. faecalis* proved to be an efficient strategy, and may be a key enzyme for further research.

Research articles 2 and 3 are contributions to the fields of transcriptomics and ^{13}C-MFA with research article 2 using ^{13}C-MFA to help characterise hydroxy acid producing variants of the TAM strain. This work compares the effects of heterologous production of lactate, malate, and 3HP, and shows that these inserted pathways cause widespread changes in the host metabolic flux networks. Research article 3 uses ^{13}C-MFA and transcriptomic data to generate predictions for gene deletions that are carried out *in vivo*. This article shows how -omic data can successfully be used to increase 3HP titres.

1.3 References

Abdel-Rahman, M.A., Tashiro, Y., Zendo, T., Shibata, K., and Sonomoto, K. (2011). Isolation and characterisation of lactic acid bacterium for effective fermentation of cellobiose into optically pure homo L-(+)-lactic acid. Appl. Microbiol. Biotechnol. *89*, 1039–1049.

Alam, M.T., Zelezniak, A., Mulleder, M., Shliaha, P., Schwarz, R., Capuano, F., Vowinckel, J., Radmanesfahar, E., Kruger, A., Calvani, E., et al. (2016). The metabolic background is a global player in Saccharomyces gene expression epistasis. Nat. Microbiol. *1*, 15030.

Andreessen, B., Lange, A.B., Robenek, H., and Steinbuchel, A. (2010). Conversion of glycerol to poly(3-hydroxypropionate) in recombinant Escherichia coli. Appl. Environ. Microbiol. *76*, 622–626.

Antoniewicz, M.R. (2015). Methods and advances in metabolic flux analysis: a mini-review. J. Ind. Microbiol. Biotechnol. *42*, 317–325.

Aung, H.W., Henry, S.A., and Walker, L.P. (2013). Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. Ind. Biotechnol. (New Rochelle. N. Y). *9*, 215–228.

Ayyoob, M., Lee, D.H., Kim, J.H., Nam, S.W., and Kim, Y.J. (2017). Synthesis of poly(glycolic acids) via solution polycondensation and investigation of their thermal degradation behaviors. Fibers Polym. *18*, 407–415.

Baart, G.J.E., and Martens, D.E. (2012). Genome-scale metabolic models: reconstruction and analysis. Methods Mol. Biol. *799*, 107–126.

Baek, S.-H., Kwon, E.Y., Bae, S.-J., Cho, B.-R., Kim, S.-Y., and Hahn, J.-S. (2017). Improvement of d-Lactic Acid Production in Saccharomyces cerevisiae Under Acidic Conditions by Evolutionary and Rational Metabolic Engineering. Biotechnol. J. *12*.

Baganz, F., Hayes, A., Farquhar, R., Butler, P.R., Gardner, D.C., and Oliver, S.G. (1998). Quantitative analysis of yeast gene function using competition experiments in continuous culture. Yeast *14*, 1417–1427.

Bai, Y., Wang, S., Zhong, H., Yang, Q., Zhang, F., Zhuang, Z., Yuan, J., Nie, X., and Wang, S. (2015). Integrative analyses reveal transcriptome-proteome correlation in biological pathways and secondary

metabolism clusters in A. flavus in response to temperature. Sci. Rep. *5*, 14582.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature *533*, 452–454.

Battat, E., Peleg, Y., Bercovitz, A., Rokem, J.S., and Goldberg, I. (1991). Optimization of L-malic acid production by Aspergillus flavus in a stirred fermentor. Biotechnol. Bioeng. *37*, 1108–1116.

Begley, C.G., and Ellis, L.M. (2012). Raise standards for preclinical cancer research. Nature *483*, 531.

Benchaita, T. (2013). Greenhouse Gas Emissions from New Petrochemical Plants Background Information Paper for the Elaboration of Technical Notes and Guidelines for IDB Projects (Inter-American Development Bank).

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology *151*, 2551–2561.

Borodina, I., and Nielsen, J. (2014). Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals. Biotechnol. J. *9*, 609–620.

Borodina, I., Kildegaard, K.R., Jensen, N.B., Blicher, T.H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M.J., Rosenstand, I., et al. (2015). Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. Metab. Eng. *27*, 57–64.

Botstein, D., and Fink, G.R. (2011). Yeast: An Experimental Organism for 21st Century Biology. Genetics *189*, 695–704.

Botstein, D., Chervitz, S.A., and Cherry, J.M. (1997). Yeast as a Model Organism. Science *277*, 1259–1260.

Branduardi, P., Sauer, M., De Gioia, L., Zampella, G., Valli, M., Mattanovich, D., and Porro, D. (2006). Lactate production yield from engineered yeasts is dependent from the host background, the lactate dehydrogenase source and the lactate export. Microb. Cell Fact. *5*, 4.

Brown, S.H., Bashkirova, L., Berka, R., Chandler, T., Doty, T., McCall, K., McCulloch, M., McFarland, S., Thompson, S., Yaver, D., et al. (2013). Metabolic engineering of Aspergillus oryzae NRRL 3488 for increased production of L-malic acid. Appl. Microbiol. Biotechnol. *97*, 8903–8912.

Burgard, A.P., Pharkya, P., and Maranas, C.D. (2003). Optknock: a bilevel programming framework for

identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

Çakar, Z.P., Sauer, U., and Bailey, J.E. (1999). Metabolic engineering of yeast: the perils of auxotrophic hosts. Biotechnol. Lett. *21*, 611–616.

Cardoso, J.G.R., Jensen, K., Lieven, C., Lærke Hansen, A.S., Galkina, S., Beber, M., Özdemir, E., Herrgård, M.J., Redestig, H., and Sonnenschein, N. (2018). Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories. ACS Synth. Biol. *7*, 1163–1166.

Chan, K.-M., Liu, Y.-T., Ma, C.-H., Jayaram, M., and Sau, S. (2013). The 2 micron plasmid of Saccharomyces cerevisiae: a miniaturized selfish genome with optimized functional competence. Plasmid *70*, 2–17.

Chen, N., Wang, J., Zhao, Y., and Deng, Y. (2018). Metabolic engineering of Saccharomyces cerevisiae for efficient production of glucaric acid at high titer. Microb. Cell Fact. *17*, 67.

Chen, Y., Zhang, Y., Siewers, V., and Nielsen, J. (2015). Ach1 is involved in shuttling mitochondrial acetyl units for cytosolic C2 provision in Saccharomyces cerevisiae lacking pyruvate decarboxylase. FEMS Yeast Res. *15*, fov015-fov015.

Chiulan, I., Frone, N.A., Brandabur, C., and Panaitescu, M.D. (2018). Recent Advances in 3D Printing of Aliphatic Polyesters. Bioeng. *5*.

Cho, S., Shin, J., and Cho, B.-K. (2018). Applications of CRISPR/Cas System to Bacterial Metabolic Engineering. Int. J. Mol. Sci. *19*.

Codon, A.C., Benitez, T., and Korhola, M. (1998). Chromosomal polymorphism and adaptation to specific industrial environments of Saccharomyces strains. Appl. Microbiol. Biotechnol. *49*, 154–163.

Colombié, S., Dequin, S., and Sablayrolles, J.M. (2003). Control of lactate production by Saccharomyces cerevisiae expressing a bacterial LDH gene. Enzyme Microb. Technol. *33*, 38–46.

Curran, K.A., Leavitt, J.M., Karim, A.S., and Alper, H.S. (2013). Metabolic engineering of muconic acid production in Saccharomyces cerevisiae. Metab. Eng. *15*, 55–66.

Curran, K.A., Crook, N.C., Karim, A.S., Gupta, A., Wagman, A.M., and Alper, H.S. (2014). Design of synthetic yeast promoters via tuning of nucleosome architecture. Nat. Commun. *5*, 4002.

Dai, Z., Liu, Y., Guo, J., Huang, L., and Zhang, X. (2015). Yeast synthetic biology for high-value metabolites. FEMS Yeast Res. *15*, 1–11.

Dalwadi, M.P., King, J.R., and Minton, N.P. (2017). Multi-timescale analysis of a metabolic network in synthetic biology: a kinetic model for 3-hydroxypropionic acid production via beta-alanine. J. Math. Biol.

De Deken, R.H. (1966). The Crabtree effect: a regulatory system in yeast. J. Gen. Microbiol. *44*, 149–156.

Delepine, B., Libis, V., Carbonell, P., and Faulon, J.-L. (2016). SensiPath: computer-aided design of sensing-enabling metabolic pathways. Nucleic Acids Res. *44*, W226-31.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res *41*.

Van Dien, S. (2013). From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. Curr. Opin. Biotechnol. *24*, 1061–1068.

Do, J.H., and Choi, D.-K. (2007). Aflatoxins: Detection, toxicity, and biosynthesis. Biotechnol. Bioprocess Eng. *12*, 585–593.

Drewke, C., Thielen, J., and Ciriacy, M. (1990). Ethanol formation in adh0 mutants reveals the existence of a novel acetaldehyde-reducing activity in Saccharomyces cerevisiae. J. Bacteriol. *172*, 3909–3917.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Endy, D. (2005). Foundations for engineering biology. Nature *438*, 449–453.

Fehér, T., Planson, A.-G., Carbonell, P., Fernández-Castané, A., Grigoras, I., Dariy, E., Perret, A., and Faulon, J.-L. (2014). Validation of RetroPath, a computer-aided design tool for metabolic pathway engineering. Biotechnol. J. *9*, 1446–1457.

Fernie, A.R., Carrari, F., and Sweetlove, L.J. (2004). Respiratory metabolism: glycolysis, the TCA cycle and

mitochondrial electron transport. Curr. Opin. Plant Biol. *7*, 254–261.

Fiume, M.M. (2017). Alpha Hydroxy Acids. Int. J. Toxicol. *36*, 15S–21S.

Fiume, Z. (2001). Final report on the safety assessment of Malic Acid and Sodium Malate. Int. J. Toxicol. *20 Suppl 1*, 47—55.

Fletcher, E., Krivoruchko, A., and Nielsen, J. (2015). Industrial systems biology and its impact on synthetic biology of yeast cell factories. Biotechnol. Bioeng. *113*, 1164–1170.

Flikweert, M.T., van der Zanden, L., Janssen, W.M.T.M., Yde Steensma, H., van Dijken, J.P., and Pronk, J.T. (1996). Pyruvate decarboxylase: An indispensable enzyme for growth of Saccharomyces cerevisiae on glucose. Yeast *12*, 247–257.

Flikweert, M.T., van Dijken, J.P., and Pronk, J.T. (1997). Metabolic responses of pyruvate decarboxylase-negative Saccharomyces cerevisiae to glucose excess. Appl. Environ. Microbiol. *63*, 3399–3404.

Flikweert, M.T., de Swaaf, M., van Dijken, J.P., and Pronk, J.T. (1999). Growth requirements of pyruvate-decarboxylase-negative Saccharomyces cerevisiae. FEMS Microbiol. Lett. *174*, 73–79.

Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., and Palsson, B.O. (2005). In silico design and adaptive evolution of Escherichia coli for production of lactic acid. Biotechnol. Bioeng. *91*, 643–648.

Förster, J., Famili, I., Fu, P., Palsson, B.Ø., and Nielsen, J. (2003). Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network. Genome Res. *13*, 244–253.

França, T.F.A., and Monserrat, J.M. (2018). Reproducibility crisis in science or unrealistic expectations? EMBO Rep.

Fredenberg, S., Wahlgren, M., Reslow, M., and Axelsson, A. (2011). The mechanisms of drug release in poly(lactic-co-glycolic acid)-based drug delivery systems—A review. Int. J. Pharm. *415*, 34–52.

Fumi, O., Toshio, F., Takehisa, N., Nobuki, T., Toru, O., Osamu, K., Toshihiro, K., and Satoshi, Y. (2009). Efficient production of L-lactic acid by Crabtree-negative yeast Candida boidinii. Yeast *26*, 485–496.

Gancedo, J.M. (1998). Yeast carbon catabolite repression. Microbiol. Mol. Biol. Rev. *62*, 334–361.

Gao, C., Ma, C., and Xu, P. (2011). Biotechnological routes based on lactic acid production from biomass. Biotechnol. Adv. *29*, 930–939.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67.

Gentile, P., Chiono, V., Carmagnola, I., and Hatton, P. V (2014). An Overview of Poly(lactic-co-glycolic) Acid (PLGA)-Based Biomaterials for Bone Tissue Engineering. Int. J. Mol. Sci. *15*, 3640–3659.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., et al. (2008). Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. Science (80-.). *319*, 1215 LP-1220.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. *2*, 31–34.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., et al. (1996). Life with 6000 genes. Science *274*, 546,563-567.

Gonzalez, A., Larroy, C., Biosca, J.A., and Arino, J. (2008). Use of the TRP1 auxotrophic marker for gene disruption and phenotypic analysis in yeast: a note of warning. FEMS Yeast Res. *8*, 2–5.

Guan, N., Du, B., Li, J., Shin, H.-D., Chen, R.R., Du, G., Chen, J., and Liu, L. (2018). Comparative genomics and transcriptomics analysis-guided metabolic engineering of Propionibacterium acidipropionici for improved propionic acid production. Biotechnol. Bioeng. *115*, 483–494.

Guaragnella, N., Palermo, V., Galli, A., Moro, L., Mazzoni, C., and Giannattasio, S. (2014). The expanding role of yeast in cancer research and diagnosis: insights into the function of the oncosuppressors p53 and BRCA1/2. FEMS Yeast Res. *14*, 2–16.

Gueldener, U., Heinisch, J., Koehler, G.J., Voss, D., and Hegemann, J.H. (2002). A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Res. *30*, e23.

Guldener, U., Heck, S., Fielder, T., Beinhauer, J., and Hegemann, J.H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res. *24*, 2519–2524.

Gunge, N. (1983). Yeast DNA plasmids. Annu. Rev. Microbiol. *37*, 253–276.

Guo, L., Zhang, F., Zhang, C., Hu, G., Gao, C., Chen, X., and Liu, L. (2018). Enhancement of malate production through engineering of the periplasmic rTCA pathway in Escherichia coli. Biotechnol. Bioeng. *115*, 1571–1580.

He, Y.-C., Xu, J.-H., Su, J.-H., and Zhou, L. (2010). Bioproduction of Glycolic Acid from Glycolonitrile with a New Bacterial Isolate of Alcaligenes sp. ECU0401. Appl. Biochem. Biotechnol. *160*, 1428–1440.

Hedbacker, K., and Carlson, M. (2008). SNF1/AMPK pathways in yeast. Front. Biosci. *13*, 2408–2420.

Hofvendahl, K., and Hahn–Hägerdal, B. (2000). Factors affecting the fermentative lactic acid production from renewable resources1. Enzyme Microb. Technol. *26*, 87–107.

Holo, H. (1989). Chloroflexus aurantiacus secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO2 and acetate. Arch. Microbiol. *151*, 252–256.

Hong, K.-K., and Nielsen, J. (2012). Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. Cell. Mol. Life Sci. *69*, 2671–2690.

Ilmén, M., Koivuranta, K., Ruohonen, L., Suominen, P., and Penttilä, M. Efficient Production of l-Lactic Acid from Xylose by Pichia stipitis. Appl. Environ. Microbiol. *73*, 117–123.

Ishida, N., Saitoh, S., Tokuhiro, K., Nagamori, E., Matsuyama, T., Kitamoto, K., and Takahashi, H. (2005). Efficient Production of l-Lactic Acid by Metabolically Engineered Saccharomyces cerevisiae with a Genome-Integrated l-Lactate Dehydrogenase Gene. Appl. Environ. Microbiol. *71*, 1964–1970.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. *169*, 5429–5433.

Jakočiūnas, T., Rajkumar, A.S., Zhang, J., Arsovska, D., Rodriguez, A., and Jendresen, C.B. (2015). CasEMBLR: Cas9-Facilitated Multiloci Genomic Integration of in Vivo Assembled DNA Parts in Saccharomyces cerevisiae. ACS Synth Biol.

Jamshidian, M., Tehrany, E.A., Imran, M., Jacquot, M., and Desobry, S. (2010). Poly-Lactic Acid: Production, applications, nanocomposites, and release studies. Compr. Rev. Food Sci. Food Saf. *9*, 552–571.

Jensen, M.K., and Keasling, J.D. (2015). Recent applications of synthetic biology tools for yeast metabolic engineering. FEMS Yeast Res. *15*, 1–10.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., and Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. *14*, 238–248.

Jessop-Fabre, M.M., Jakociunas, T., Stovicek, V., Dai, Z., Jensen, M.K., Keasling, J.D., and Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. Biotechnol. J. *11*, 1110–1117.

Jiang, X., Meng, X., and Xian, M. (2009). Biosynthetic pathways for 3-hydroxypropionic acid production. Appl. Microbiol. Biotechnol. *82*, 995–1003.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Juturu, V., and Wu, J.C. (2016). Microbial production of lactic acid: the latest development. Crit. Rev. Biotechnol. *36*, 967–977.

Kaniak, A., Xue, Z., Macool, D., Kim, J.-H., and Johnston, M. (2004). Regulatory network connecting two glucose signal transduction pathways in Saccharomyces cerevisiae. Eukaryot. Cell *3*, 221–231.

Karim, A.S., Curran, K.A., and Alper, H.S. (2013). Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications. FEMS Yeast Res. *13*, 10.1111/1567-1364.12016.

Karlheinz, M. (2000). Hydroxycarboxylic Acids, Aliphatic. In Ullmann's Encyclopedia of Industrial Chemistry, (American Cancer Society), p.

Kildegaard, K.R., Hallstrom, B.M., Blicher, T.H., Sonnenschein, N., Jensen, N.B., Sherstyk, S., Harrison, S.J., Maury, J., Herrgard, M.J., Juncker, A.S., et al. (2014). Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. Metab. Eng. *26*, 57–66.

Kildegaard, K.R., Jensen, N.B., Schneider, K., Czarnotta, E., Ozdemir, E., Klein, T., Maury, J., Ebert, B.E., Christensen, H.B., Chen, Y., et al. (2016). Engineering and systems-level analysis of Saccharomyces cerevisiae for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway. Microb. Cell Fact. *15*, 53.

Kim, H.U., Kim, T.Y., and Lee, S.Y. (2008). Metabolic flux analysis and metabolic engineering of microorganisms. Mol. Biosyst. *4*, 113–120.

Kim, W.J., Kim, H.U., and Lee, S.Y. (2017). Current state and applications of microbial genome-scale metabolic models. Curr. Opin. Syst. Biol. *2*, 10–18.

Kornhauser, A., Coelho, S.G., and Hearing, V.J. (2010). Applications of hydroxy acids: classification, mechanisms, and photoactivity. Clin. Cosmet. Investig. Dermatology CCID *3*, 135–142.

Kouprina, N., and Larionov, V. (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. Chromosoma *125*, 621–632.

Kozak, B.U., van Rossum, H.M., Luttik, M.A.H., Akeroyd, M., Benjamin, K.R., Wu, L., de Vries, S., Daran, J.-M., Pronk, J.T., and van Maris, A.J.A. (2014). Engineering acetyl coenzyme A supply: functional expression of a bacterial pyruvate dehydrogenase complex in the cytosol of Saccharomyces cerevisiae. MBio *5*, e01696-14.

Krivoruchko, A., Siewers, V., and Nielsen, J. (2011). Opportunities for yeast metabolic engineering: lessons from synthetic biology. Biotechnol *6*.

Kumar, V., Ashok, S., and Park, S. (2013). Recent advances in biological production of 3-hydroxypropionic acid. Biotechnol. Adv. *31*, 945–961.

Kunkee, R.E. (1968). Malo-Lactic Fermentation. W.W.B.T.-A. in A.M. Umbreit, ed. (Academic Press), pp. 235–279.

Lee, J.Y., Kang, C.D., Lee, S.H., Park, Y.K., and Cho, K.M. (2015). Engineering cellular redox balance in Saccharomyces cerevisiae for improved production of L-lactic acid. Biotechnol. Bioeng. *112*, 751–758.

Lee, S.Y., Park, J.M., and Kim, T.Y. (2011). Application of metabolic flux analysis in metabolic engineering. Methods Enzymol. *498*, 67–93.

Leng, G., and Song, K. (2016). Watch out for your TRP1 marker: the effect of TRP1 gene on the growth at

high and low temperatures in budding yeast. FEMS Microbiol. Lett. *363*.

Lerman, J.A., Hyduke, D.R., Latif, H., Portnoy, V.A., Lewis, N.E., Orth, J.D., Schrimpe-Rutledge, A.C., Smith, R.D., Adkins, J.N., Zengler, K., et al. (2012). In silico method for modelling metabolism and gene product expression at genome scale. Nat. Commun. *3*, 929.

Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., et al. (2010). Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. *6*, 390.

Li, M., and Borodina, I. (2015). Application of synthetic biology for production of chemicals in yeast Saccharomyces cerevisiae. FEMS Yeast Res. *15*, 1–12.

Li, M., Kildegaard, K.R., Chen, Y., Rodriguez, A., Borodina, I., and Nielsen, J. (2015). De novo production of resveratrol from glucose or ethanol by engineered Saccharomyces cerevisiae. Metab. Eng. *32*, 1–11.

Li, Y., Wang, X., Ge, X., and Tian, P. (2016). High Production of 3-Hydroxypropionic Acid in Klebsiella pneumoniae by Systematic Optimization of Glycerol Metabolism. Sci. Rep. *6*, 26932.

Li, Z.-J., Hong, P.-H., Da, Y.-Y., Li, L.-K., and Stephanopoulos, G. (2018). Metabolic engineering of Escherichia coli for the production of L-malate from xylose. Metab. Eng. *48*, 25–32.

Lian, J., HamediRad, M., and Zhao, H. (2018). Advancing Metabolic Engineering of Saccharomyces cerevisiae Using the CRISPR/Cas System. Biotechnol. J. e1700601.

Liu, J., Li, J., Shin, H.-D., Du, G., Chen, J., and Liu, L. (2017a). Metabolic engineering of Aspergillus oryzae for efficient production of l-malate directly from corn starch. J. Biotechnol. *262*, 40–46.

Liu, J., Li, J., Shin, H., Du, G., Chen, J., and Liu, L. (2017b). Biological production of l-malate: recent advances and future prospects. World J. Microbiol. Biotechnol. *34*, 6.

Lui, C., and Lievense, J. (2005). Lactic acid producing yeast - US Patent application 20050112737.

Lynen, F. (1959). Participation of acyl—coa in carbon chain biosynthesis. J. Cell. Comp. Physiol. *54*, 33–49.

Machado, D., and Herrgård, M.J. (2015). Co-evolution of strain design methods based on flux balance and elementary mode analysis. Metab. Eng. Commun. *2*, 85–92.

Mack, M., Gompel-Klein, P., Haase, E., Hietkamp, J., Ruhland, A., and Brendel, M. (1988). Genetic characterization of hyperresistance to formaldehyde and 4-nitroquinoline-N-oxide in the yeast Saccharomyces cerevisiae. Mol. Gen. Genet. *211*, 260–265.

Mans, R., Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., and Daran-lapujade, P. (2015). CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res *15*.

Maris, A.J.A. va., Konings, W.N., Dijken, J.P. va., and Pronk, J.T. (2004). Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. Metab. Eng. *6*, 245–255.

van Maris, A.J.A., Geertman, J.-M.A., Vermeulen, A., Groothuizen, M.K., Winkler, A.A., Piper, M.D.W., van Dijken, J.P., and Pronk, J.T. (2004a). Directed Evolution of Pyruvate Decarboxylase-Negative Saccharomyces cerevisiae, Yielding a C2-Independent, Glucose-Tolerant, and Pyruvate-Hyperproducing Yeast. Appl. Environ. Microbiol. *70*, 159–166.

van Maris, A.J.A., Winkler, A.A., Porro, D., van Dijken, J.P., and Pronk, J.T. (2004b). Homofermentative Lactate Production Cannot Sustain Anaerobic Growth of Engineered Saccharomyces cerevisiae: Possible Consequence of Energy-Dependent Lactate Export. Appl. Environ. Microbiol. *70*, 2898–2905.

Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. Nature *526*, 55.

Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., and Porro, D. (2012). Recombinant Protein Production in Yeasts BT - Recombinant Gene Expression. A. Lorence, ed. (Totowa, NJ: Humana Press), pp. 329–358.

Maury, J., Germann, S.M., Baallal Jacobsen, S.A., Jensen, N.B., Kildegaard, K.R., Herrgård, M.J., Schneider, K., Koza, A., Forster, J., Nielsen, J., et al. (2016). EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in Saccharomyces cerevisiae. PLoS One *11*, e0150394.

Maury, J., Kannan, S., Jensen, N.B., Öberg, F.K., Kildegaard, K.R., Forster, J., Nielsen, J., Workman, C.T., and Borodina, I. (2018). Glucose-Dependent Promoters for Dynamic Regulation of Metabolic Pathways . Front. Bioeng. Biotechnol. *6*, 63.

Mazumdar, S., Clomburg, J.M., and Gonzalez, R. Escherichia coli Strains Engineered for Homofermentative Production of d-Lactic Acid from Glycerol . Appl. Environ. Microbiol. *76*, 4327–4336.

McAlister-Henn, L., Steffan, J.S., Minard, K.I., and Anderson, S.L. (1995). Expression and Function of a Mislocalized Form of Peroxisomal Malate Dehydrogenase (MDH3) in Yeast. J. Biol. Chem. *270*, 21220–21225.

McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., Nuñez, A., Butrym, E.D., Richards, M.P., Wang, C., et al. (2004). Fermented beverages of pre- and proto-historic China. Proc. Natl. Acad. Sci. U. S. A. *101*, 17593–17598.

Meussen, B.J., de Graaff, L.H., Sanders, J.P.M., and Weusthuis, R.A. (2012). Metabolic engineering of Rhizopus oryzae for the production of platform chemicals. Appl. Microbiol. Biotechnol. *94*, 875–886.

Miano, J.M., Zhu, Q.M., and Lowenstein, C.J. (2016). A CRISPR Path to Engineering New Genetic Mouse Models for Cardiovascular Research. Arterioscler. Thromb. Vasc. Biol. *36*, 1058 LP-1075.

Mojica, F.J.M., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. *60*, 174–182.

Molenaar, D., van Berlo, R., de Ridder, D., and Teusink, B. (2009). Shifts in growth strategies reflect tradeoffs in cellular economics. Mol. Syst. Biol. *5*.

Morales, M., Ataman, M., Badr, S., Linster, S., Kourlimpinis, I., Papadokonstantakis, S., Hatzimanikatis, V., and Hungerbühler, K. (2016). Sustainability assessment of succinic acid production technologies from biomass using metabolic engineering. Energy Environ. Sci. *9*, 2794–2805.

Murad, H., Shamban, A.T., and Premo, P.S. (1995). The use of glycolic acid as a peeling agent. Dermatol. Clin. *13*, 285–307.

Nagamori, E., Shimizu, K., Fujita, H., Tokuhiro, K., Ishida, N., and Takahashi, H. (2013). Metabolic flux analysis of genetically engineered Saccharomyces cerevisiae that produces lactate under micro-aerobic conditions. Bioprocess Biosyst. Eng. *36*, 1261–1265.

Nanchen, A., Fuhrer, T., and Sauer, U. (2007). Determination of Metabolic Flux Ratios From 13C-Experiments and Gas Chromatography-Mass Spectrometry Data BT - Metabolomics: Methods and Protocols. W. Weckwerth, ed. (Totowa, NJ: Humana Press), pp. 177–197.

NEST (2005). Synthetic Biology: Applying Engineering to Biology: Report of a NEST High Level Expert Group Luxembourg: Office for Official Publications of the European Communities.

Nevoigt, E. (2008). Progress in Metabolic Engineering of Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. *72*, 379–412.

Nielsen, J. (2014). Synthetic Biology for Engineering Acetyl Coenzyme A Metabolism in Yeast. MBio *5*, e02153-14.

Nielsen, J. (2015). Yeast cell factories on the horizon. Science (80-.). *349*, 1050 LP-1051.

Nielsen, J., and Keasling, J.D. (2017). Engineering Cellular Metabolism. Cell *164*, 1185–1197.

Nielsen, J., Larsson, C., van Maris, A., and Pronk, J. (2013). Metabolic engineering of yeast for production of fuels and chemicals. Curr. Opin. Biotechnol. *24*, 398–404.

Orth, J.D., Thiele, I., and Palsson, B.Ø. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Oud, B., Flores, C.-L., Gancedo, C., Zhang, X., Trueheart, J., Daran, J.-M., Pronk, J.T., and van Maris, A.J.A. (2012). An internal deletion in MTH1 enables growth on glucose of pyruvate-decarboxylase negative, non-fermentative Saccharomyces cerevisiae. Microb. Cell Fact. *11*, 131.

Ozsolak, F., and Milos, P.M. (2010). RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. *12*, 87.

Papagianni, M. (2012). Metabolic engineering of lactic acid bacteria for the production of industrially important compounds. Comput. Struct. Biotechnol. J. *3*, e201210003.

Pasteur, L. (1995). Mémoire sur la fermentation appelée lactique (Extrait par l'auteur). Mol. Med. *1*, 599–601.

Peter, J., De Chiara, M., Friedrich, A., Yue, J.-X., Pflieger, D., Bergström, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., et al. (2018). Genome evolution across 1,011 Saccharomyces cerevisiae isolates. Nature *556*, 339–344.

Pfeiffer, T., and Morley, A. (2014). An evolutionary perspective on the Crabtree effect. Front. Mol. Biosci. *1*, 17.

Pines, O., Even-Ram, S., Elnathan, N., Battat, E., Aharonov, O., Gibson, D., and Goldberg, I. (1996). The cytosolic pathway of l-malic acid synthesis in Saccharomyces cerevisiae: the role of fumarase. Appl. Microbiol. Biotechnol. *46*, 393–399.

Polumienko, A.L., Grigor'eva, S.P., Lushnikov, A.A., and Domaradskij, I. V (1986). Yeast centromeric plasmids as shuttle vectors between Escherichia coli, Bacillus subtilis and Saccharomyces cerevisiae. Biochem. Biophys. Res. Commun. *135*, 915–921.

Pronk, J.T., Yde Steensma, H., and Van Dijken, J.P. (1996). Pyruvate metabolism in Saccharomyces cerevisiae. Yeast *12*, 1607–1633.

Quinn, J.Y., Cox III, R.S., Adler, A., Beal, J., Bhatia, S., Cai, Y., Chen, J., Clancy, K., Galdzicki, M., Hillson, N.J., et al. (2015). SBOL Visual: A Graphical Language for Genetic Designs. PLOS Biol. *13*, e1002310.

Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie *117*, 119–128.

Rathin, D., and Michael, H. (2006). Lactic acid: recent advances in products, processes and technologies — a review. J. Chem. Technol. Biotechnol. *81*, 1119–1129.

Reed, A.M., and Gilding, D.K. (1981). Biodegradable polymers for use in surgery — poly(glycolic)/poly(lactic acid) homo and copolymers: 2. In vitro degradation. Polymer (Guildf). *22*, 494–498.

Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R.A., Tong, G.J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N.J., et al. (2017). A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. Nucleic Acids Res. *45*, 496–508.

Runguphan, W., and Keasling, J.D. (2014). Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. Metab Eng *21*.

Saitoh, S., Ishida, N., Onishi, T., Tokuhiro, K., Nagamori, E., Kitamoto, K., and Takahashi, H. (2005). Genetically Engineered Wine Yeast Produces a High Concentration of l-Lactic Acid of Extremely High Optical Purity. Appl. Environ. Microbiol. *71*, 2789–2792.

Sánchez, B.J., Zhang, C., Nilsson, A., Lahtvee, P., Kerkhoven, E.J., and Nielsen, J. (2017). Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. Mol. Syst. Biol. *13*.

Sarazin, P., Roy, X., and Favis, B.D. (2004). Controlled preparation and properties of porous poly(L-lactide) obtained from a co-continuous blend of two biodegradable polymers. Biomaterials *25*, 5965–5978.

Sauer, B. (1987). Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol. Cell. Biol. *7*, 2087–2096.

Sauer, U. (2004). High-throughput phenomics: experimental methods for mapping fluxomes. Curr. Opin. Biotechnol. *15*, 58–63.

Sauer, M., Porro, D., Mattanovich, D., and Branduardi, P. (2010). 16 years research on lactic acid production with yeast - ready for the market? Biotechnol. Genet. Eng. Rev. *27*, 229–256.

Sene, M., Gilmore, I., and Janssen, J.-T. (2017). Metrology is key to reproducing results. Nature *547*, 397–399.

Sherman, F. (2002). Getting started with yeast. Methods Enzymol. *350*, 3–41.

Shi, S., Chen, Y., Siewers, V., and Nielsen, J. (2014). Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. MBio. *5*.

Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.-. K., and Keasling, J.D. (2007). Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metab Eng *9*.

Shirra, M.K., Patton-Vogt, J., Ulrich, A., Liuta-Tehlivets, O., Kohlwein, S.D., Henry, S.A., and Arndt, K.M. (2001). Inhibition of Acetyl Coenzyme A Carboxylase Activity Restores Expression of the INO1 Gene in a snf1Mutant Strain of Saccharomyces cerevisiae . Mol. Cell. Biol. *21*, 5710–5722.

Da Silva, N.A., and Srikrishnan, S. (2012). Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. FEMS Yeast Res. *12*, 197–214.

Skjoedt, M.L., Snoek, T., Kildegaard, K.R., Arsovska, D., Eichenberger, M., Goedecke, T.J., Rajkumar, A.S., Zhang, J., Kristensen, M., Lehka, B.J., et al. (2016). Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. Nat. Chem. Biol. *12*, 951.

Snoep, J.L., Westphal, A.H., Benen, J.A.E., de Mattos, M.J., Neijssel, O.M., and De Kok, A. (1992). Isolation and characterisation of the pyruvate dehydrogenase complex of anaerobically grown Enterococcus faecalis NCTC 775. Eur. J. Biochem. *203*, 245–250.

Sobolov, M., and Smiley, K.L. (1960). Metabolism of glycerol by an acrolein-forming lactobacillus. J. Bacteriol. *79*, 261–266.

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. Proc. Natl. Acad. Sci.

Stephanopoulos, G. (1999). Metabolic fluxes and metabolic engineering. Metab. Eng. *1*, 1–11.

Stinchcomb, D.T., Mann, C., and Davis, R.W. (1982). Centromeric DNA from Saccharomyces cerevisiae. J. Mol. Biol. *158*, 157–190.

Stovicek, V., Borodina, I., and Forster, J. (2015). CRISPR–Cas system enables fast and simple genome editing of industrial Saccharomyces cerevisiae strains. Metab. Eng. Commun. *2*, 13–22.

Stovicek, V., Holkenbrink, C., and Borodina, I. (2017). CRISPR/Cas system for yeast genome engineering: advances and applications. FEMS Yeast Res. *17*.

Tabita, F.R. (2009). The hydroxypropionate pathway of CO2 fixation: Fait accompli. Proc. Natl. Acad. Sci. U. S. A. *106*, 21015–21016.

Talarico, T.L., and Dobrogosz, W.J. (1990). Purification and Characterization of Glycerol Dehydratase from Lactobacillus reuteri. Appl. Environ. Microbiol. *56*, 1195–1197.

Tan, J., Abdel-Rahman, M.A., and Sonomoto, K. (2018). Biorefinery-Based Lactic Acid Fermentation: Microbial Production of Pure Monomer Product BT - Synthesis, Structure and Properties of Poly(lactic acid). M.L. Di Lorenzo, and R. Androsch, eds. (Cham: Springer International Publishing), pp. 27–66.

Tehlivets, O., Scheuringer, K., and Kohlwein, S.D. (2007). Fatty acid synthesis and elongation in yeast. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1771*, 255–270.

Thompson-Jaeger, S., Francois, J., Gaughran, J.P., and Tatchell, K. (1991). Deletion of SNF1 affects the nutrient response of yeast and resembles mutations which activate the adenylate cyclase pathway. Genetics *129*, 697–706.

Trumbly, R.J. (1992). Glucose repression in the yeast Saccharomyces cerevisiae. Mol. Microbiol. *6*, 15–21.

Tsao, G.T., Cao, N.J., Du, J., and Gong, C.S. (1999). Production of multifunctional organic acids from renewable resources. Adv. Biochem. Eng. Biotechnol. *65*, 243–280.

Upadhyaya, B.P., DeVeaux, L.C., and Christopher, L.P. (2014). Metabolic engineering as a tool for enhanced

lactic acid production. Trends Biotechnol. *32*, 637–644.

Vanegas, K.G., Lehka, B.J., and Mortensen, U.H. (2017). SWITCH: a dynamic CRISPR tool for genome engineering and metabolic pathway control for cell factory construction in Saccharomyces cerevisiae. Microb Cell Fact *16*.

Vickers, C.E., Bydder, S.F., Zhou, Y., and Nielsen, L.K. (2013). Dual gene expression cassette vectors with antibiotic selection markers for engineering in Saccharomyces cerevisiae. Microb. Cell Fact. *12*, 96.

Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793–1808.

Wach, A., Brachat, A., Alberti-Segui, C., Rebischung, C., and Philippsen, P. (1997). Heterologous HIS3 marker and GFP reporter modules for PCR-targeting in Saccharomyces cerevisiae. Yeast *13*, 1065–1075.

Wakil, S.J., Stoops, J.K., and Joshi, V.C. (1983). Fatty Acid Synthesis and Its Regulation. Annu. Rev. Biochem. *52*, 537–579.

Wang, L., Dash, S., Ng, C.Y., and Maranas, C.D. (2017). A review of computational tools for design and reconstruction of metabolic pathways. Synth. Syst. Biotechnol. *2*, 243–252.

Wehner, E.P., Rao, E., and Brendel, M. (1993). Molecular structure and genetic regulation of SFA, a gene responsible for resistance to formaldehyde in Saccharomyces cerevisiae, and characterization of its protein product. Mol. Gen. Genet. MGG *237*, 351–358.

Werpy, T., and Petersen, G. (2004). Top Value Added Chemicals from Biomass: Volume I -- Results of Screening for Potential Candidates from Sugars and Synthesis Gas (United States).

Wilson, W.A., Hawley, S.A., and Hardie, D.G. (1996). Glucose repression/derepression in budding yeast: SNF1 protein kinase is activated by phosphorylation under derepressing conditions, and this correlates with a high AMP:ATP ratio. Curr. Biol. *6*, 1426–1434.

Yang, T.H. (2013). 13C-Based Metabolic Flux Analysis: Fundamentals and Practice BT - Systems Metabolic Engineering: Methods and Protocols. H.S. Alper, ed. (Totowa, NJ: Humana Press), pp. 297–334.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., et al. (2011). Metabolic engineering of Escherichia coli for direct production of

1,4-butanediol. Nat. Chem. Biol. *7*, 445.

Yin, X., Li, J., Shin, H., Du, G., Liu, L., and Chen, J. (2015). Metabolic engineering in the biotechnological production of organic acids in the tricarboxylic acid cycle of microorganisms: Advances and prospects. Biotechnol. Adv. *33*, 830–841.

Yun, J.-S., Wee, Y.-J., and Ryu, H.-W. (2003). Production of optically pure l(+)-lactic acid from various carbohydrates by batch fermentation of Enterococcus faecalis RKY1. Enzyme Microb. Technol. *33*, 416–423.

Yung, T.W., Jonnalagadda, S., Balagurunathan, B., and Zhao, H. (2016). Transcriptomic Analysis of 3-Hydroxypropanoic Acid Stress in Escherichia coli. Appl. Biochem. Biotechnol. *178*, 527–543.

Zaman, S., Lippman, S.I., Zhao, X., and Broach, J.R. (2008). How Saccharomyces Responds to Nutrients. Annu. Rev. Genet. *42*, 27–81.

Zamboni, N., Fendt, S.-M., Rühl, M., and Sauer, U. (2009). 13C-based metabolic flux analysis. Nat. Protoc. *4*, 878.

Zelle, R.M., de Hulster, E., van Winden, W.A., de Waard, P., Dijkema, C., Winkler, A.A., Geertman, J.-M.A., van Dijken, J.P., Pronk, J.T., and van Maris, A.J.A. (2008). Malic Acid Production by Saccharomyces cerevisiae: Engineering of Pyruvate Carboxylation, Oxaloacetate Reduction, and Malate Export . Appl. Environ. Microbiol. *74*, 2766–2777.

Zelle, R.M., de Hulster, E., Kloezen, W., Pronk, J.T., and van Maris, A.J.A. (2010). Key Process Conditions for Production of C4 Dicarboxylic Acids in Bioreactor Batch Cultures of an Engineered Saccharomyces cerevisiae Strain . Appl. Environ. Microbiol. *76*, 744–750.

Zhang, X., Wang, X., Shanmugam, K.T., and Ingram, L.O. (2015). L-malate production by metabolically engineered escherichia coli.

Research Articles

Research Article 1

EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into *Saccharomyces cerevisiae* via CRISPR-Cas9

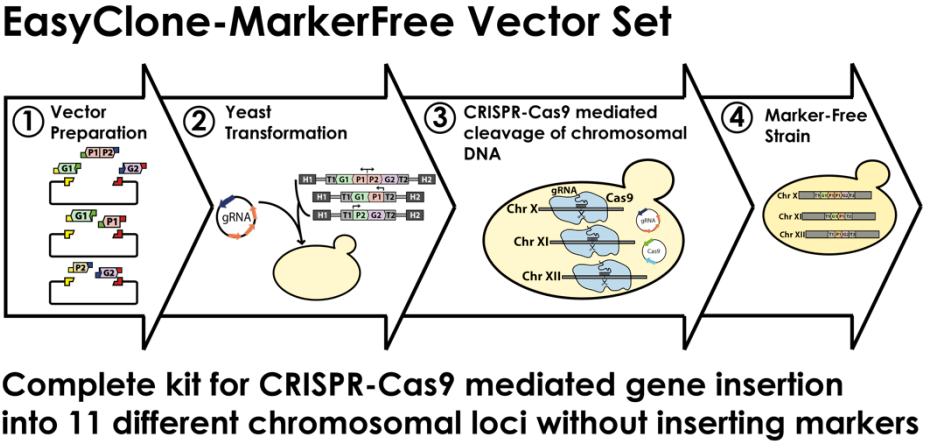
Mathew M Jessop-Fabre[1,*], Tadas Jakočiūnas[1,*], Vratislav Stovicek[1], Zongjie Dai[2,3], Michael K Jensen[1], Jay D Keasling[1,4,5,6], and Irina Borodina[1]. *Biotechnology Journal*. 2016, 11(8): 1110-1117.

[1] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
[2] The Novo Nordisk Foundation Center for Biosustainability, Chalmers University of Technology, Gothenburg, Sweden
[3] Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden
[4] Joint BioEnergy Institute, Emeryville, CA, USA
[5] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[6] Department of Chemical and Biomolecular Engineering & Department of Bioengineering University of California, Berkeley, CA, USA

[*] These authors contributed equally to this work

Abstract

Saccharomyces cerevisiae is an established industrial host for production of recombinant proteins, fuels and chemicals. To enable stable integration of multiple marker-free overexpression cassettes in the genome of *S. cerevisiae*, we have developed a vector toolkit EasyClone-MarkerFree. The integration of linearized expression cassettes into defined genomic loci is facilitated by CRISPR/Cas9. Cas9 is recruited to the chromosomal location by specific guide RNAs (gRNAs) expressed from a set of gRNA helper vectors. Using our genome engineering vector suite, single and triple insertions are obtained with 90–100% and 60–70% targeting efficiency, respectively. We demonstrate application of the vector toolkit by constructing a haploid laboratory strain (CEN.PK113-7D) and a diploid industrial strain (Ethanol Red) for production of 3-hydroxypropionic acid, where we tested three different acetyl-CoA supply strategies, requiring overexpression of three to six genes each. Among the tested strategies was a bacterial cytosolic pyruvate dehydrogenase complex, which was integrated into the genome in a single transformation. The publicly available EasyClone-MarkerFree vector suite allows for facile and highly standardized genome engineering, and should be of particular interest to researchers working on yeast chassis with limited markers available.



Graphical abstract

1 Introduction

The yeast *S. cerevisiae* is widely applied in industrial biotechnology for production of fuels, chemicals, and pharmaceutical ingredients and it is also used for fundamental research as a eukaryotic model organism [1-3]. Genetic manipulation of *S. cerevisiae* is greatly facilitated by its efficient inherent homologous recombination (HR) machinery for DNA double-strand break (DSB) repair [4]. As a result, many genome engineering tools have been developed that take advantage of HR for targeted integration of heterologous DNA into the yeast genome [5, 6]. Chromosomal gene integration has several benefits over plasmid-based approaches, such as increased strain stability, better control of gene expression level and lower population heterogeneity [5]. However, though the HR machinery has a high fidelity, selection is still necessary to identify the colonies that have been successfully engineered. This has the fundamental drawback that selection markers need to be recycled if multiple genome edits are needed. Also, both dominant and auxotrophic markers have been reported to have an effect on cell physiology [7, 8]. Moreover, antibiotic-based selection markers are undesirable in industrial strains due to the risk of spreading drug resistance. A way to improve homology-directed chromosomal integration of heterologous gene fragments is by using DNA endonucleases for targeted DSB as this has been shown to increase recombination events by approximately three orders of magnitude [9-12]. Recently, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) Type II system from bacteria has been adopted for genetic engineering of yeast [13]. This system consists of two components: guide RNA (gRNA) and CRISPR-associated endonuclease (Cas9). Upon expression in the cell, the gRNA/Cas9 complex is recruited to DNA by base-pairing of the gRNA recognition sequence (typically 20 bp) and the complementary DNA stand, after which the Cas9 introduces a DSB break. The target DNA sequence must be immediately followed by a protospacer adjacent motif (PAM) (i.e. NGG for *Streptococcus pyogenes*-derived Cas9); Cas9 makes a blunt cut three nucleotides upstream of the PAM site [14]. There have been many recently published examples of CRISPR-Cas9-based tools for yeast genome

engineering [15-21] that are reviewed in Jakočiūnas et al. [22]. The work presented here builds on the previously reported method EasyClone for single-step targeted integration of multiple expression cassettes [5, 6]. The EasyClone method allows targeted genomic integration of up to three vectors in a single transformation event using auxotrophic or dominant selection markers. These target sites are located in the intergenic regions and are interspaced by essential genes to ensure that the integrated DNA fragments are not at risk of being removed by HR [23]. Moreover it has been confirmed that integration into these sites ensures a high-level of expression and does not interfere with cellular growth [23]. The selection markers can be removed via Cre-LoxP-mediated recombination, which requires one extra transformation with CreA-vector, cultivation to induce the CreA expression and finally screening of the resulting clones on selection plates to confirm the loss of the marker; the marker removal procedure has a turnaround time of about five days. Once the selection markers are eliminated, the strain is ready for the next integration event or for other genetic modifications. Marker removal, as well as being time consuming, can also cause genome instability [24, 25].

Here we present the EasyClone-MarkerFree method, which takes advantage of CRISPR-Cas9 for introduction of double-strand DNA breaks in the defined integration sites, which causes very efficient integration of expression cassettes and thus bypasses the need for selection. The gRNA helper vectors, expressing gRNA molecules that recruit Cas9 to the particular chromosomal locations, can be easily removed from the strain by growth on non-selective medium, after which the strain is ready for the next modification. This shortens the turnaround time to only one or two days. Additional advantages are: simplified experimental design for cloning, because the choice of selection markers does not need to be taken into consideration; elimination of the effect of the selection markers on strain physiology; and moreover a possibility to iterate gene integrations with other edits guided by CRISPR/Cas9, e.g. gene deletions. The use of predefined and characterized chromosomal target sites dispenses with the need for the experimenter to identify suitable regions in which to integrate their genes of interest. Furthermore, we demonstrate the

utility of the method by engineering 3-hydroxypropionic acid producing strains with improved precursor supply strategies.

2 Materials and methods

2.1 Strains

Two different strains of *S. cerevisiae* were used: the haploid laboratory strain CEN.PK113-7D (*MATa URA3 HIS3 LEU2 TRP1 MAL2-8c SUC2*; obtained from Peter Kötter, Johann Wolfgang Goethe University Frankfurt, Germany), and the diploid industrial Ethanol Red strain (*MAT*a/α; obtained from Fermentis, A Lesaffre division). *E. coli* strain DH5α alpha was used to clone, propagate, and store the plasmids.

2.2 Construction of plasmids

The codon-optimized genes, encoding the pyruvate dehydrogenase complex and lipoate-protein ligase from *Enterococcus faecalis*, were ordered from GenScript (sequences can be found in Supporting information, Table S5). The genes encoding ATP-dependent citrate lyase and the mitochondrial citrate transport protein, were amplified from the genomic DNA of *Yarrowia lipolytica* DSM-8218, obtained from the DSMZ collection (www.dsmz.de). All of the primers, biobricks and plasmids constructed and used in this study can be found in Supporting information, Tables S1–S3.

The EasyClone-MarkerFree vectors were created by amplifying the EasyClone 2.0 vectors [6] with primers that were designed to attach to either side of the selection markers, creating a fragment that no longer contained the marker. These fragments were then ligated to form the marker-less vectors. Seven of the resulting vectors (named "Intermediate vectors" in Table S3) contained PAM sites in the integration regions, which were removed by site-directed mutagenesis using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturers' protocol.

gRNA cassettes targeting particular integration loci (chromosomal coordinates can be found in Supporting information, Table S4) were ordered as double-stranded gene blocks from IDT DNA. These cassettes were amplified using primers 10525(TJOS-62 [P1F]) and 10529(TJOS-65 [P1R]) and USER-cloned into pCfB2926 (pTAJAK-71) [15] to give single gRNA helper vectors. For construction of triple gRNA helper vectors, three gRNA cassettes were amplified using three primer pairs (10525(TJOS-62 [P1F]) and 10530(TJOS-66 [P2R]) for the first, 10526(TJOS-63 [P2F]) and 10531(TJOS-67 [P3R]) for the second, and 10527(TJOS-64 [P3F]) and 10529(TJOS-65 [P1R]) for the third gRNA cassette) and cloned into pCfB2926 (p-TAJAK-71) [15]. Single gRNA helper vectors for Ethanol Red were constructed by PCR amplification of the template plasmid pCfB3041 using primers indicated in Supporting information, Table S1 as described in [17]. All of the cloning steps for creating gRNA helper vectors and EasyClone-MarkerFree vectors were performed in *E. coli*. Correct cloning was confirmed by Sanger sequencing.

For expression of the Cas9 gene we used an episomal vector pCfB2312 with CEN-ARS replicon and KanMX resistance marker [17].

The EasyClone-MarkerFree vectors for expression of fluorescent protein or 3HP pathway genes were cloned as described in [5, 26]. The vectors were linearized with *Not*I, the integration fragment (part of the expression vector without *E. coli* ori and AmpR) was gel-purified and transformed, along with a gRNA helper vector, into yeast carrying the Cas9 plasmid (pCfB2312) via the lithium acetate method [27]. After the heat shock the cells were recovered for two hours in YPD medium and then plated on YPD agar containing 200 mg/L G418 and 100 mg/L nourseothricin. For yeast transformations with a single vector we routinely use 500 ng of the linear integration fragment along with 500 ng of the relevant gRNA helper plasmid. For yeast transformations with three vectors we use 1 μg of linear integration fragments and 1 μg of triple gRNA helper plasmid. Correct integration

of the vectors into the genome was verified by colony PCR using primers listed in Supporting information, Table S1.

2.3 Media and growth conditions

The *E. coli* strains containing the plasmids were grown in Luria–Bertani (LB) media supplemented with 100 μg/mL ampicillin. Yeast strains were grown in liquid YPD media, or on YPD agar plates with 20 g/L glucose (Sigma-Aldrich). When needed the medium was supplemented with antibiotics G418 (200 mg/L) for the selection of the Cas9 plasmid, and/or nourseothricin (100 mg/L) for selection of the gRNA plasmid. Synthetic complete medium with 20 g/L glucose was prepared using drop-out powder from Sigma-Aldrich.

The strains were screened for production of 3HP as following. Single colonies were grown in 500 μL synthetic complete media in 96 deep-well plates with air-penetrable lids (EnzyScreen, Germany) at 30°C with 250 rpm shaking overnight. The overnight cultures were then inoculated into 50 mL of either mineral media (MM) with 60 g/L glucose [5], or a simulated fed-batch media (feed-in-time, FIT) in 500 mL-shake flasks with baffles to a starting OD600 of 0.1. The feed-in-time (FIT) medium M-Sc.syn-1000 synthetic FIT was purchased from M2P labs GmbH (Germany). All the media for 3HP cultivations were supplemented with 500 ng/mL of DL-α-lipoic acid (Sigma Aldrich). The strains were cultivated in MM or FIT media at 30°C with 250 rpm for 168 h. Samples were taken every 12 h (every 1–2 h during the exponential phase in MM), with the optical density OD_{600} being measured on a 20-fold diluted culture broth. Samples were then centrifuged at 3000 $\times g$ for 30 min and the supernatant was stored at −20°C until HPLC analysis. Cultivations were carried out in triplicate.

2.4 Analytical procedures

Strains expressing *gfp* were analyzed using a BioTek Synergy MX microplate reader. Excitation wavelength was set at 485 nm, and emission was read at 530 nm. Optical density was measured at 600 nm. Concentration of 3-hydroxypropionic acid in the broth was measured by HPLC, using a method previously reported in [28].

2.5 Integration Stability

Strains that had been transformed with *gfp*-containing vectors were grown in 500 μL of YPD in 96 deep-well plates with air-penetrable lids at 30°C with 250 rpm shaking overnight. The following morning, 50 μL of each culture was passed to 500 μL of fresh YPD. This was repeated for a total of five passages. At the end of the passages, the culture was plated onto YPD agar, and the fluorescent colonies were counted.

2.6 Determination of gene copy number by qPCR

Quantitative PCR was performed on the Ethanol Red strain containing the genes from *Enterococcus faecalis*. Primers were designed using the PrimerQuest® tool (https://eu.idtdna.com/Primerquest/Home/Index). Details of the primers used can be found in Supporting information, Table S1. Genomic DNA was prepared using the ZR Fungal/Bacterial DNA MiniPrep™ kit (Zymo Research) according to the manufacturer's protocol. Each reaction contained 10 μL of 2xSYBR Green qPCR MasterMix (Life Technologies), 1 μL of each primer, 0.1 μL of diluted reference dye, and 7.7 μL of genomic DNA. Reactions were carried out in triplicate, with 5 pg of gDNA present in each reaction. *ALG9* was used as a control, as well as a no template control where the gDNA was replaced with nuclease-free PCR grade water. The experiment was carried out using a Stratagene Mx3005P qPCR system with the following settings; 10 min pre-incubation at 95°C, 40 cycles of denaturation at 95°C for 20 s and annealing and elongation at 60°C for 22 s, and the final step of 95°C for 1 min, 55°C for 30 s, 95°C for 30 s. The Stratagene

MxPro software was used to quantify the original amount of DNA present for each gene, which could then be related to the copy number.

3 Results and discussion

To enable rapid iterative cycles of metabolic engineering, we developed a new vector toolkit EasyClone-MarkerFree, which allows integration of genes into specific sites of the yeast chromosome without integration of a selection marker. This simplifies strain design and reduces the turnaround time in comparison to the previously published EasyClone methods [5, 6]. The method employs CRISPR-Cas9 to introduce one or several DSBs into the host chromosome, ensuring efficient integration. For this system to be active in *S. cerevisiae*, it is necessary to have a Cas9 expression plasmid, as well as a gRNA helper vector. The DSB is repaired by the host cell using the native HR machinery, which incorporates the EasyClone-MarkerFree vector(s) into the target loci via double cross-over.

3.1 Construction of EasyClone-MarkerFree vector toolkit

Construction of the EasyClone vector set required that 11 gRNA helper vectors were built, one to target each of the EasyClone chromosomal integration sites in the *S. cerevisiae* CEN.PK strain. Furthermore, three triple gRNA helper vectors were built that can target three sites simultaneously. The EasyClone 2.0 vectors were modified to remove the selection markers and, where necessary, to mutate the PAM sites in the integration sites. The vectors contain a USER cloning site surrounded by bi-directional terminators. Genes and promoters can be cloned into the vectors using USER cloning as described below or other methods, such as Gibson assembly, in-fusion, MoClo, etc. [29].

3.2 Workflow

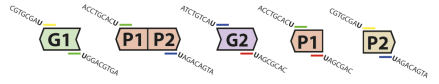
The suggested workflow is illustrated on Fig. 1. For USER cloning, the EasyClone-MarkerFree vectors are prepared by linearization with *Asi*SI and nicking with Nb.*Bsm*I in order to obtain vector backbones with sticky ends. The treated vector backbones are stored at −20°C. Compatible biobricks, encoding genes and promoters, are prepared by PCR amplification using oligos with a uracil base, 7–10 bases downstream of the 5' end of the primers (Fig. 1a). We routinely use the overhangs as specified in [5] for all of our genes and promoters, which simplifies standardization and exchange of biobricks between scientists in our laboratory. We store the library of ready-to-use genes and promoters at −20°C.

To clone promoters and genes into the expression vector, the *Asi*SI/Nb.*Bsm*I-treated vector backbones are mixed with the desired promoter(s) and gene(s) and treated with USER enzyme. The USER enzyme excises uracils in the overhangs of genes and promoters, thus producing sticky ends compatible with the EasyClone vectors (Fig. 1b). The USER reaction mix is transformed into *E. coli*, which ligates the expression vector (Fig. 1c). The plasmids can then be isolated, confirmed by sequencing and linearized with *Not*I (or another suitable) restriction enzyme to produce linear integration fragments containing expression cassettes flanked with ~500-bp-long integration regions (Fig. 1d). The linear integration fragment(s) and the corresponding helper vector are then transformed into the yeast host cell already carrying a Cas9 expression plasmid (Fig. 1e). The transformants are selected on medium with G418 and nourseothricin to select for the Cas9 and gRNA helper vector, respectively. If further CRISPR-Cas9 mediated engineering rounds are desired, the gRNA helper vector is removed by cultivating the cells in liquid or solid medium containing G418 to maintain the selection for Cas9 (Fig. 1f). gRNA helper vectors are 2μ-based and according to our experience are easily lost by growing the cells overnight in liquid medium with G418 only. The resulting culture is diluted and plated on

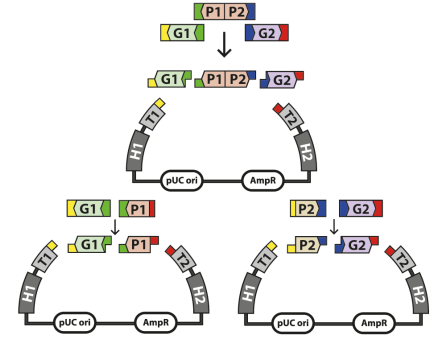
G418 plates to obtain single colonies. The loss of the gRNA helper vector is confirmed by replica plating onto nourseothricin agar plates. The strains are then ready for the next round of Cas9-mediated gene insertions or other manipulations (e.g. knock-outs, etc.). If no further genetic modifications are desired, the strains can be cultivated on non-selective medium to eliminate both Cas9 and gRNA helper vectors.

Figure 1 (Over page). An overview of the EasyClone-MarkerFree method. The BioBricks encoding genes and promoters are generated by PCR-amplification with uracil-containing primers (a). The BioBricks are assembled with the integration vectors via USER cloning (b) and the reaction mixture is transformed into E. coli (c). The resulting plasmids are isolated from E. coli and confirmed by sequencing, then they are linearized and transformed into Cas9-expressing S. cerevisiae along with a helper gRNA vector, which causes double-stranded breaks at the designated integration sites (d–e). Yeast cells are selected on plates with G418 and nourseothricin and correct integration of the vector(s) is confirmed by PCR. The helper gRNA vector can be eliminated by growth in the absence of nourseothricin selection and the resulting strain, which still expresses Cas9, can be further engineered. Once the desired genetic modifications are accomplished the Cas9 vector is removed by growth on non-selective medium and the final strain, not containing selection markers, is obtained (f). G, gene; P, promoter; T, terminator; H, homology region. A detailed protocol is provided in Supporting information.

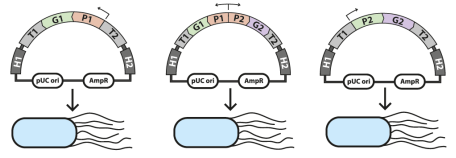
a) Amplification of promoters and genes with overhangs to create BioBricks



b) Promoters and genes are cloned into AsiSI and Nb.BsmI treated EasyClone-MarkerFree vector backbones

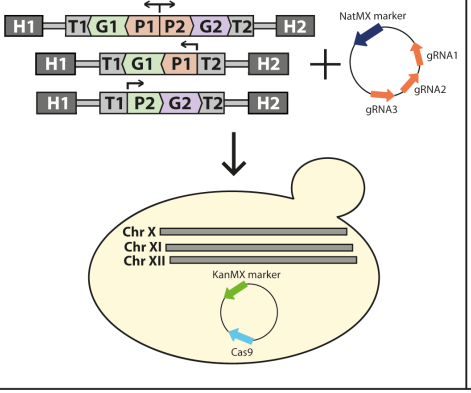


c) The vectors are transformed into *E. coli*

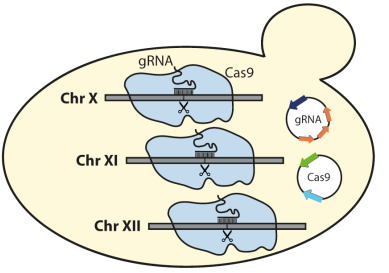


d) The expression vectors are linearized

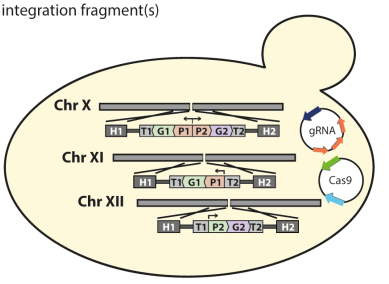
e) The linear integration fragments along with gRNA helper vector are transformed into *S. cerevisiae* expressing Cas9. The transformants are selected on plates with G418 and nourseothricin.



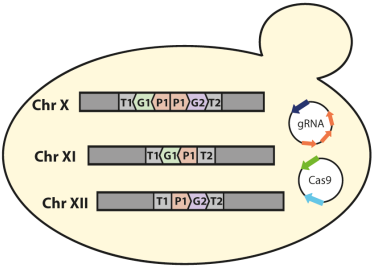
e(i). The gRNA produced will guide the Cas9 endonuclease to the chromosome and introduce a double stranded break



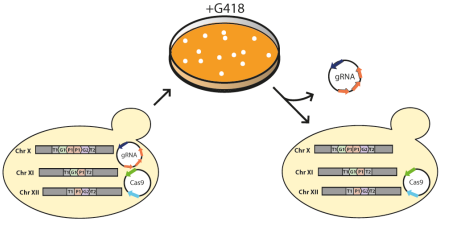
e(ii). The native homologous recombination machinery of *S. cerevisiae* repaires the DSB with the introduced linear integration fragment(s)



e(iii). The linear fragments are stably integrated into the designated loci on the chromosomes



f) The gRNA helper vector is removed by growing the cells on non-selective medium. The strain is now ready for a new round of genetic modifications



3.3 Validation of the vector toolkit

To verify the integration efficiency and gene expression levels at each site, *gfp* was cloned into each of the EasyClone vectors and transformed into the CEN.PK113-7D strain. The integrations were tested by colony PCR for 20 randomly picked clones, and 95–100% correct integrations were obtained for all the integration sites (Fig. 2). This is comparable with the previous versions of EasyClone, where the targeting efficiency was 80–100% [5, 6]. However, we should note that wrong clones in the EasyClone-MarkerFree system were not fluorescent and hence they did not contain the designed integration cassettes in their genomes, while for the EasyClone system with selection markers, the integration cassette would always be integrated in the genome, however sometimes it would not be targeted to the designed integration site. Further, the transformed colonies were each checked for fluorescence, with 99% of the colonies showing fluorescence (84–559 colonies were obtained and tested per transformation). Fluorescence levels showed some smaller variations for different integration sites (Fig. 2) as we also observed previously [6]. Another observation was that among those clones with an integrated fragment, all clones were fluorescent, which differs from the results we obtained with CRISPR-Cas9-mediated insertion of expression fragments with short homology arms (40 bp), where only half of the clones were fluorescent [17]. We suggest that the high fidelity of correct integration is due to the long homology arms of ~0.5 kb present in the EasyClone vectors.

We further tested the stability of the integrations of the *gfp* gene in each of the sites. After five passages, the population retained 100% of fluorescent colonies, reinforcing the notion that these sites are stable due to their location in-between essential genes.

The simultaneous targeting of three integration sites, assisted with triple gRNA helper vectors, resulted in 60–70% targeting efficiencies, which is also comparable with previously reported efficiencies for CRISPR-Cas9 methods for multi-loci integration [16, 18]. We

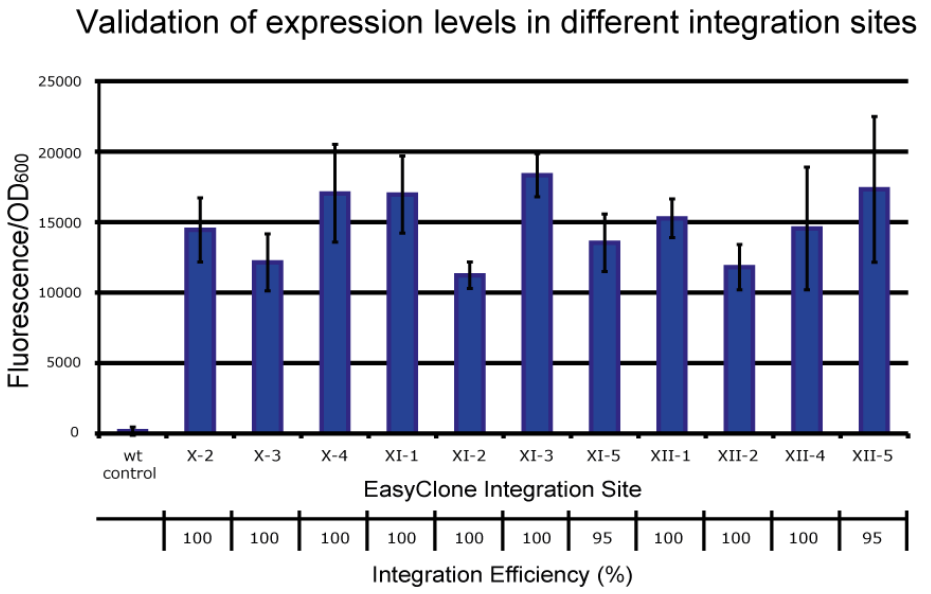


Figure 2. Validation of the EasyClone-MarkerFree vector toolkit. The laboratory strain, CEN.PK 117-7D, was transformed with the EasyClone-MarkerFree vectors carrying gfp gene under control of a constitutive PTEF1 promoter. Specific fluorescence was measured to evaluate the expression level at each of the eleven integration sites. The error bars show standard deviation between nine or ten replicates. The percentage of clones with correct integration is shown for each site n = 20.

previously observed for the EasyClone system with selection markers that targeting efficiency into the same chromosome was only 44%, however the efficiency is higher for simultaneous integration into different chromosomes (unpublished results). We believe that simultaneous targeting of three sites that are closely positioned on the same chromosome, may lead to genome instability. Therefore to achieve the best targeting efficiencies, triple gRNA helper vectors were designed to target distinct chromosomes.

3.4 Application of EasyClone-MarkerFree vector toolkit for engineering of precursor supply for 3-hydroxypropionic acid production

As a test case, a pathway to produce 3HP was inserted into two different *S. cerevisiae* strains, CEN.PK113-7D and Ethanol Red, and the strains were further engineered for improved acetyl-CoA supply. The basic pathway strains were constructed by inserting a modified acetyl-CoA carboxylase *ACC1*^{S659A,S1157A} and a malonyl-CoA reductase *MCR* from *Chloroflexus aurantiacus*. Three different strategies to boost acetyl-CoA production in the cytosol were compared (Fig. 3A). One of the strategies relied on expression of a functional bacterial pyruvate dehydrogenase complex from *Enterococcus faecalis* in the cytoplasm, which required expression of genes encoding subunits E1α, E1β, E2, and E3 of PDH complex, as well as two genes involved in lipoylation of E2 [30]. Another strategy was based on overexpression of the pyruvate dehydrogenase bypass, namely pyruvate decarboxylase *PDC1*, aldehyde dehydrogenase *ALD6* and acetyl-CoA synthase from *Salmonella enterica ACSSE* [31]. The third strategy encompassed overexpression of ATP-dependent citrate lyase, consisting of two subunits *ACL1* and *ACL2*, and a mitochondrial citrate transporter protein from *Y. lipolytica*. The genes required for different strategies were cloned into three EasyClone-MarkerFree plasmids and transformed into yeast strains expressing basic 3HP pathway along with the triple gRNA helper vector (pCfB3052 for CEN.PK and pCfB4668 for Ethanol Red).

Overexpression of the cytoplasmic PDH complex proved to be the most successful strategy in CEN.PK, with an improvement in 3HP final titer of 19% over the basic strain in mineral medium, and 95% in the simulated fed-batch medium (Fig. 3). The PDH bypass showed no improvement. The ACL strategy performed worse than the basic pathway in MM, but interestingly showed an improvement of 60% in FIT (Fig. 3B). Growth profiles can be found in Supporting information, Fig. S1.

In the diploid industrial strain (Ethanol Red), the PDH complex also performed well, with an improvement of 23% over the basic strain (Supporting information, Fig. S2). These

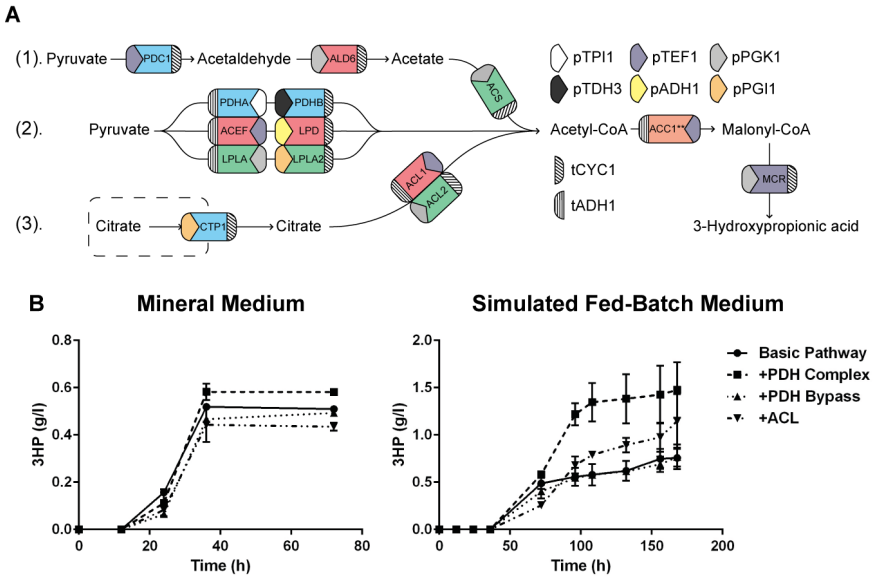


Figure 3. Effect of acetyl-CoA supply strategies on the production of 3-hydroxypropionic acid. (A) Three different pathways taken to produce acetyl-CoA in the cytosol; 1. The PDH bypass, 2. The PDH complex of E. faecalis, 3. The ATP-dependent citrate lyase strategy from Y. lipolytica. Colored arrows denote which promoter was used, and the direction of the arrow indicates in which direction the gene is transcribed. A left facing arrow signifies that the gene is translated on the anti-sense strand and vice versa. The box at the end of the gene represents which terminator sequence is used. (B) Production of 3-hydroxypropionic acid over time in the engineered laboratory strain in mineral or simulated fed-batch medium. The cultivations were carried out in triplicate, the error bars show standard deviation.

results confirm that the bacterial PDH complex from *Enterococcus faecalis* is a robust and efficient way to improve cytosolic acetyl-CoA levels in yeast. We confirmed that EasyClone-MarkerFree method integrates the expression cassettes on both homologous chromosomes, when applied in diploid strains. We performed qPCR on the diploid industrial strain (Ethanol Red) carrying the genes for the PDH complex. All of the six genes were shown to have a copy number of 2. The PDH bypass seems to have performed poorly, contradicting previous works. This may be due to the low activity of the 3HP biosynthetic genes in this work, as we only integrated *ACC1* and *MCR* in single copies. In other studies, where we expressed *ACC1* and *MCR* from either a high-copy episomal

vector or integrated them in multiple copies into the genome, the overexpression of PDH bypass did have a positive effect on the 3HP production [32-34].

4 Concluding remarks

The EasyClone-MarkerFree vector toolkit can be used to simultaneously introduce one to three integration cassettes into the genome of *S. cerevisiae* without the use of selection markers. Using standardized BioBricks, the integration cassettes can be constructed for overexpression of one or two genes per integration site. In this study we successfully integrated up to six genes in a single transformation with 60–70% targeting efficiency. The system is well suited for strain construction via multiple iterative metabolic engineering cycles. We have shown that it performs well in both the haploid laboratory CEN.PK strain, and also in the diploid industrial Ethanol Red strain. The EasyClone-MarkerFree vector toolkit can be obtained from AddGene.

Acknowledgement

The work was funded by the Novo Nordisk Foundation. VS and IB acknowledge funding for the BioREFINE-2G project by the European Commission in the 7th Framework Programme (Project no. FP7-613771).

The authors declare no financial or commercial conflict of interest.

Supplementary Information

Table S1. List of primers used in this study

ID	NAME	SEQUENCE (5'->3')	APPLICATION
Primers used for construction of EasyClone-MarkerFree plasmids and gRNA helper plasmids			
401	ID401pIntFwdU	ACCCAAUTCGCCCTATAGTGAGTCG	For amplification of the EasyClone vectors minus the markers
402	ID402pintRevU	ACGCGAUCTTCGAGCGTCCCAAAACC	For amplification of the EasyClone vectors minus the markers
10277	gRNArev	GATCATTTATCTTTCACTGCGGAGAAG	universal primer for amplification of gRNA cassette carrying plasmid
11150	X-2_gRNAerFW	CTCTCGAAGTGGTCACGTGTGTTTTAGAGCTAGAA	For amplification of X-2 targeting gRNA helper vector for ethanol red, used with 10277
11151	X-4_gRNAerFW	CGCCATTCAAGAGTAGCAACGTTTTAGAGCTAGAA	For amplification of X-2 targeting gRNA helper vector for ethanol red, used with 10277
11152	XI-2_gRNAerFW	TTGATCAGTTGATCAGTTGAGTTTTAGAGCTAGAA	For amplification of X-2 targeting gRNA helper vector for ethanol red, used with 10277
11153	XI-5_gRNAerFW	TGAGAATAATGTTGTAAAACGTTTTAGAGCTAGAA	For amplification of X-2 targeting gRNA helper vector for ethanol red, used with 10277
10525	TJOS-62 (P1F)	CGTGCGAUAGGGAACAAAAGCTGGAGCT	Amplify 1st gRNA expression sequence for triple gRNA vectors
10526	TJOS-63 (P2F)	AGTGCAGGUAGGGAACAAAAGCTGGAGCT	Amplify 2nd gRNA expression sequence for triple gRNA vectors
10527	TJOS-64 (P3F)	ATCTGTCAUAGGGAACAAAAGCTGGAGCT	Amplify 3rd gRNA expression sequence for triple gRNA vectors
10529	TJOS-65 (P1R)	CACGCGAUTAACTAATTACATGACTCGA	Amplify 3rd gRNA expression sequence for triple gRNA vectors
10530	TJOS-66 (P2R)	ACCTGCACUTAACTAATTACATGACTCGA	Amplify 1st gRNA expression sequence for triple gRNA vectors
10531	TJOS-67 (P3R)	ATGACAGAUTAACTAATTACATGACTCGA	Amplify 2nd gRNA expression sequence for triple gRNA vectors
Primers used for PAM site removal by site-directed mutagenesis			
11399	TJOS-120	ATTAATGCCTCAGCACTAGT	Universal primer for PAM site removal in EasyClone vectors
11354	TJOS-89F	CCACTTTTCAATGAAACGGA	Primer to remove PAM site in pCfB2900
11395	TJOS-116	ACATGGGAAGATTCGCTTTT	Primer to remove PAM site in pCfB2901
11396	TJOS-117	AGTTTCTTGGCATTGGCAAT	Primer to remove PAM site in pCfB2902
11397	TJOS-118	CTATTGGCTGCTTCATAGTA	Primer to remove PAM site in pCfB2905
11398	TJOS-119	AGTTTACTCAATTCTTGAAG	Primer to remove PAM site in pCfB2908
11400	TJOS-121	CGTGAATCAACTGCACATAC	Primer to remove PAM site in pCfB2906
11355	TJOS-90F	CGACTCTCTCGAAATTTTTC	Primer to remove PAM site in pCfB2907

Primers used for amplification of gene and promoter BioBricks			
6	PTEF1_rv	CACGCGAUGCACACACCATAGCTTC	For amplification of the TEF1 promoter
7	PPGK1_fw	CGTGCGAUGGAAGTACCTTCAAAGA	For amplification of the PGK1 promoter
53	ACC1m_fw	CGT GCG AUT CAT TTC AAA GTC TTC AAC AAT TT	For amplification of ACC1** gene
177	CaMCR_rv_NEW	CACGCGAUTCAGACTGTAATGGCTCTACCTC	For amplification of MCR gene
312	GFPopt_rv	CACGCGAU TCA TTTGTAGAGCTCATCCATGC	For amplification of GFP gene
645	ACSse_U1_rv (ID645)	CGTGCGAUTCATGATGGCATAGCAATAG	For amplification of ACS gene
739	ald6_U2_rv (ID739)	CACGCGAUTCACAACTTAATTCTGACAGCTTTTAC	For amplification of ALD6 gene
842	YlAcly1_U1_rv (ID842)	CGTGCGAU TCA TGATCGAGTCTTGGCCTTG	For amplification of ACL1 gene
844	YlAcly2_U2_rv (ID844)	CACGCGAU TCA AACTCCGAGAGGAGTGGAAG	For amplification of ACL2 gene
846	YlCtp1_U1_rv (ID846)	CGTGCGAU TCA AAGAATCTCCATGATCTTC	For amplification of CTP1 gene
1188	pdc1_U1longer_rv (ID1188)	CGTGCGAUTCATTGCTTAGCGTTGGTAGCAGCAGTC	For amplification of PDC1 gene
1564	PTEF1->_fw (ID1564)	CGTGCGAUGCACACACCATAGCTTC	For amplification of TEF1 promoter
10549	lplA_U2_fw	ATCTGTCAUAAAACAATGAGATACGTTATCATGCAATCC	For amplification of LPLA gene
10550	lplA_U_rv	CACGCGAUTCATTAGTCAACCAACAAGTGGACG	For amplification of LPLA gene
10684	pdhB_U2_fw	ATCTGTCAUAAAACAATGGCACAAAAGACTATGATCCAAGC	For amplification of PDHB gene
10685	pdhB_U_rv	CACGCGAUTCATTAAAATTCAACAATTTCTCTGG	For amplification of PDHB gene
10686	pdhA_U1_fd	AGTGCAGGUAAAACAATGGCAAAAGCAAAGAAGC	For amplification of PDHA gene
10687	pdhA_U_rv	CGTGCGAUTTACTTTGATTCTTTGGCTTCG	For amplification of PDHA gene
10688	aceF_U2_fw	ATCTGTCAUAAAACAATGGCCTATCAATTCAAGTTGCC	For amplification of ACEF gene
10689	aceF_U_rv	CACGCGAUTCAACCTTCCATCAATAA	For amplification of ACEF gene
10690	lpd_U1_fw	AGTGCAGGUAAAACAATGGTTGTTGGTGAC	For amplification of LPD gene
10691	lpd_U_rv	CGTGCGAUTCATTAGATATGTATAGGCAAACC	For amplification of LPD gene
Sequencing primers			
891	XII-1-up-out-sq ID891	CTGGCAAGAGAACCACCAAT	Verifies UP region of chromosome site XII-1 with ID2221

893	XII-2-up-out-sq ID893	CGAAGAAGGCCTGCAATTC	Verifies UP region of chromosome site XII-2 with ID2221
897	XII-4-up-out-sq ID897	GAACTGACGTCGAAGGCTCT	Verifies UP region of chromosome site XII-4 with ID2221
899	XII-5-up-out-sq ID899	CCACCGAAGTTGATTTGCTT	Verifies UP region of chromosome site XII-5 with ID2221
901	X-2-up-out-sq ID901	TGCGACAGAAGAAAGGGAAG	Verifies UP region of chromosome site X-2 with ID2221
903	X-3-up-out-sqID903	TGACGAATCGTTAGGCACAG	Verifies UP region of chromosome site X-3 with ID2221
905	X-4-up-out-sqID905	CTCACAAAGGGACGAATCCT	Verifies UP region of chromosome site X-4 with ID2221
907	Ny-XI-1-up-sqID907	CTTAATGGGTAGTGCTTGACACG	Verifies UP region of chromosome site XI-1 with ID2221
909	XI-2-sq-fwID909	GTTTGTAGTTGGCGGTGGAG	Verifies UP region of chromosome site XI-2 with ID2221
911	XI-3-up-out-sqID911	GTGCTTGATTTGCGTCATTC	Verifies UP region of chromosome site XI-3 with ID2221
2221	JM234_ColoPCR_vec_TADH1_towards out	GTTGACACTTCTAAATAAGCGAATTTC	Verifies UP regions of chromosome sites
8418	verif ChrXI-5_up (JM269)	CTCAATGATCAAAATCCTGAATGCA	Verifies UP region of chromosome site XI-5
qPCR Primers			
14120	ALG9_qPCR_fw	CCGTTGCCATGTTGTTGTATG	Primer to use as a control for qPCR
14121	ALG9_qPCR_rv	GCCAGGAAATTGTACGCTAAAC	Primer to use as a control for qPCR
15553	(AceF qPCR) FWD	GCGACTTCTGGGTAGTTGATAA	For amplification of a short section of the E. faecalis AceF gene
15554	(AceF qPCR) REV	TAACCGCACAAGATATGAGAGAC	For amplification of a short section of the E. faecalis AceF gene
15555	(lpd qPCR) FWD	GCCTTCACAGACCCTGAATTA	For amplification of a short section of the E. faecalis lpd gene
15556	(lpd qPCR) REV	CTCTACCGTTACCAGCGAATG	For amplification of a short section of the E. faecalis lpd gene
15557	(pdhA qPCR) FWD	GTCGTCACCGGACAAAGTAT	For amplification of a short section of the E. faecalis pdhA gene
15558	(pdhA qPCR) REV	AGCCAAGGAAGCAAGAGATT	For amplification of a short section of the E. faecalis pdhA gene
15559	(lplA qPCR) FWD	CCATGCGTTATCTTGGGTAGAA	For amplification of a short section of the E. faecalis lplA gene
15560	(lplA qPCR) REV	CACCACCACCAGACAATCTT	For amplification of a short section of the E. faecalis lplA gene
15561	(lplA2 qPCR) FWD	GCACCATCAGCAACATTCATC	For amplification of a short section of the E. faecalis lplA2 gene
15562	(lplA2 qPCR) REV	CAGAAACTGGGACTGGAACTAC	For amplification of a short section of the E. faecalis lplA2 gene
15563	(pdhB qPCR) FWD	CAGAGAAGAAGTCCCAGATGAAG	For amplification of a short section of the E. faecalis pdhB gene
15564	(pdhB qPCR) REV	CATGGCACCGTAGGTTATAATAGA	For amplification of a short section of the E. faecalis pdhB gene

Table S2. List of BioBricks used in this study

ID	Name	Primers used	Template	Description
BB8	<-TEF1	5, 6	p0029	TEF1 promoter in position P1
BB157	pTEF1-ACC1_pPGK1-CaMCR	53, 177	p0298	acc1 - double mutant under control of TEF1 promoter in position 1. mcr gene from *Chloroflexus aurantiacus* under control of PGK1 promter in position 2
BB183	ctp1<-	845, 846	*Y. lipolytica* genomic DNA	Mitochondrial citrate transporter from *Y. lipolytica* in position 1
BB867	pTEF1-GFP	1564, 312	p2199	GFP protein under control of TEF1 promoter in position 2
BB903	pTDH3->	10720, 279	p2525	TDH3 promoter in position 2 to fuse to a promoter in position 1
BB904	<-pdhA	10686, 10687	Codon optimised pdhA gene string	pdhA subunit of PDH complex from *E. faecalis* in position 1
BB905	pdhB->	10684, 10685	Codon optimised pdhB gene string	pdhB subunit of PDH complex from *E. faecalis* in position 2
BB906	<-lpd	10690, 10691	Codon optimised lpd gene string	lpd subunit of PDH complex from *E. faecalis* in position 1
BB907	aceF->	10688, 10689	Codon optimised aceF gene string	aceF subunit of PDH complex from *E. faecalis* in position 2
BB908	<-pTEF1	10721, 1750	p0029	TEF1 promoter in position 1 to fuse to a promoter in position 2
BB909	pPGK1->	10722, 293	p0029	PGK1 promoter in position 1 to fuse to a promoter in position 1
BB918	<-pTPI1	10764, 10765	*S. cerevisiae* genomic DNA	TPI1 promoter in position 1 to fuse to a promoter in position 2
BB1021	<-pPGI1	10545, 10546	*S. cerevisiae* genomic DNA	PGI1 promoter in position 1 to fuse to a promoter in position 2
BB1022	lplA->	10549, 10550	Codon optimised pdhA gene string	lplA lipoylation gene of PDH complex from *E. faecalis* in position 1
BB1023	<-lplA2	10547, 10548	Codon optimised pdhA gene string	lplA2 lipoylation gene of PDH complex from *E. faecalis* in position 2
BB1127	ylacly1<-pTEF1	842, 6	p0626	acly2 gene from *Y. lipolytica* under the control of TEF1 promoter in position 1
BB1128	pPGK1->ylacly2	844, 7	p0626	acly1 gene from *Y. lipolytica* under the control of TEF1 promoter in position 2
BB1449	pdc1<-pTEF1	1188, 6	p0382	pdc1 gene under control of TEF1 promoter in position 1

BB1450	pPGK1->ALD6	739, 7	p0380	ald6 gene under control of PGK1 promoter in position 2
BB1451	ACSse<-pTEF1	645, 6	p0380	acs from *S. enterica* under control of TEF1 promoter in position 1

Table S3. List of plasmids used in this study

ID	Name	Description	Addgene reference	Source
	Cas9 expression plasmid			
pCfB2312	TEF1p-Cas9-CYC1t_kanMX	Episomal plasmid for Cas9 expression		[17]
	Intermediate vectors			
pCfB2900	X-3-intermediate for MF	Amplified from plasmid pCfB127, using primers 401 and 402 . Contains PAM site		This study
pCfB2901	X-4-intermediate for MF	Amplified from plasmid pCfB2070, using primers 401 and 402 . Contains PAM site		This study
pCfB2902	XI-1-intermediate for MF	Amplified from plasmid pCfB2328, using primers 401 and 402 . Contains PAM site		This study
pCfB2905	XI-5-intermediate for MF	Amplified from plasmid pCfB387, using primers 401 and 402 . Contains PAM site		This study
pCfB2906	XII-1-intermediate for MF	Amplified from plasmid pCfB2072, using primers 401 and 402 . Contains PAM site		This study
pCfB2907	XII-2-intermediate for MF	Amplified from plasmid pCfB120, using primers 401 and 402 . Contains PAM site		This study
pCfB2908	XII-4-intermediate for MF	Amplified from plasmid pCfB130, using primers 401 and 402 . Contains PAM site		This study
	EasyClone-MarkerFree vectors			
pCfB2899	X-2-MarkerFree backbone	Amplified from plasmid pCfB126 ,using primers 401 and 402.	73271	This study
pCfB2903	XI-2-MarkerFree	Amplified from plasmid pCfB384 ,using primers 401 and 402.	73275	This study
pCfB2904	XI-3-MarkerFree	Amplified from plasmid pCfB2118 ,using primers 401 and 402.	73276	This study
pCfB2909	XII-5-MarkerFree	Amplified from plasmid pCfB2073 ,using primers 401 and 402.	73281	This study
pCfB3034	X-3-MarkerFree	PAM site in pCfB2900 mutated with primers 11399 and 11354	73272	This study
pCfB3035	X-4-MarkerFree	PAM site in pCfB2901 mutated with primers 11399 and 11395	73273	This study
pCfB3036	XI-1-MarkerFree	PAM site in pCfB2902 mutated with primers 11399 and 11396	73274	This study
pCfB3037	XI-5-MarkerFree	PAM site in pCfB2905 mutated with primers 11399 and 11397	73277	This study

Table S3. List of plasmids used in this study

pCfB3038	XII-1-MarkerFree	PAM site in pCfB2906 mutated with primers 11399 and 11400	73278	This study
pCfB3039	XII-2-MarkerFree	PAM site in pCfB2907 mutated with primers 11399 and 11355	73279	This study
pCfB3040	XII-4-MarkerFree	PAM site in pCfB2908 mutated with primers 11399 and 11398	73280	This study
	Single target gRNA helper vectors			
pCfB3020	gRNA-X-2	gRNA sequence for targetting site X-2 USER cloned into pCfB2926	73282	This study
pCfB3041	gRNAX-3	gRNA sequence for targetting site X-3 USER cloned into pCfB2926	73283	This study
pCfB3042	gRNAX-4	gRNA sequence for targetting site X-4 USER cloned into pCfB2926	73284	This study
pCfB3043	gRNAXI-1	gRNA sequence for targetting site XI-1 USER cloned into pCfB2926	73285	This study
pCfB3044	gRNAXI-2	gRNA sequence for targetting site XI-2 USER cloned into pCfB2926	73286	This study
pCfB3045	gRNAXI-3	gRNA sequence for targetting site XI-3 USER cloned into pCfB2926	73287	This study
pCfB3046	gRNAXI-5	gRNA sequence for targetting site XI-5 USER cloned into pCfB2926	73288	This study
pCfB3047	gRNAXII-1	gRNA sequence for targetting site XII-1 USER cloned into pCfB2926	73289	This study
pCfB3048	gRNAXII-2	gRNA sequence for targetting site XII-2 USER cloned into pCfB2926	73290	This study
pCfB3049	gRNAXII-4	gRNA sequence for targetting site XII-4 USER cloned into pCfB2926	73291	This study
pCfB3050	gRNAXII-5	gRNA sequence for targetting site XII-5 USER cloned into pCfB2926	73292	This study
pCfB3588	EthanolRed gRNA X-2	Amplified with primers 10277 and 11150 , using pCfB3041 as template		This study
pCfB3589	EthanolRed gRNA X-4	Amplified with primers 10277 and 11151, using pCfB3041 as template		This study
pCfB3590	EthanolRed gRNA XI-2	Amplified with primers 10277 and 11152 , using pCfB3041 as template		This study
pCfB3591	EthanolRed gRNA XI-5	Amplified with primers 10277 and 11153 , using pCfB3041 as template		This study
	Triple target gRNA helper vectors			
pCfB3051	gRNAX-3XI-2XII-2	Three different target site cassettes amplified using primers 10525 and 10530, 10526 and 10531, 10527 and 10529 . Sections are amplified from the respective single gRNAs	73293	This study
pCfB3052	gRNAX-4XI-3XII-5	Three different target site cassettes amplified using primers 10525 and 10530, 10526 and 10531, 10527 and 10529 . Sections are amplified from the respective single gRNAs	73294	This study
pCfB3053	gRNAX-2XI-5XII-4	Three different target site cassettes amplified using primers 10525 and 10530, 10526 and 10531, 10527 and 10529 . Sections are amplified from the respective single gRNAs	73295	This study
pCfB4668	Ethanol Red p-gRNA X-4 XI-3 XII-5	Three different target site cassettes amplified using primers 10525 and 10530, 10526 and 10531, 10527 and 10529 . Sections are amplified from the respective single gRNAs		This study
	GFP Vectors			

pCfB3008	X-2-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB2899		This study
pCfB3009	X-3-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3034		This study
pCfB3010	X-4-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3035		This study
pCfB3011	XI-1-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3036		This study
pCfB3012	XI-2-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB2903		This study
pCfB3013	XI-3-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB2904		This study
pCfB3014	XI-5-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3037		This study
pCfB3015	XII-1-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3038		This study
pCfB3016	XII-2-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3039		This study
pCfB3017	XII-4-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB3040		This study
pCfB3018	XII-5-MarkerFree-GFP	Biobrick 867, USER cloned into plasmid pCfB2909		This study
	Production vectors			
pCfB3069	X-3-MF-MCR-ACC1	Biobrick 157, USER cloned into plasmid pCfB3034		This study
pCfB3472	pdhA-pdhB	Biobricks 904, 905, 903, 918, USER cloned into plasmid pCfB3035		This study
pCfB3473	aceF-lpd	Biobricks 906, 907, 908, 909, USER cloned into plasmid pCfB2904		This study
pCfB3479	lplA2-lplAXII-5	Biobricks 1022, 1023, 1021, 909, USER cloned into plasmid pCfB2909		This study
pCfB3584	X-4 ylacly1<-pTEF1	Biobrick 1127, USER cloned into plasmid pCfB3035		This study
pCfB3585	XI-3 pPGK1->ylacly2	Biobrick 1128, USER cloned into plasmid pCfB2904		This study
pCfB3587	pdc X-4	Biobrick 1449, USER cloned into plasmid pCfB3035		This study
pCfB3592	pPGK1->ALD6	Biobrick 1450, USER cloned into plasmid pCfB2904		This study
pCfB3593	ACS-pPGK1 XII-5	Biobrick 1451, USER cloned into plasmid pCfB2909		This study
pCfB3594	ylCtp1<-pTEF1 XII-5	Biobricks 183, 8 USER cloned into plasmid pCfB2909		This study

Table S4. Chromosomal coordinates of the EasyClone-MarkerFree integration sites

Integration site or Gene name	Chromosomal coordinates according to Saccharomyces genome database (http://www.yeastgenome.org/).
X-2	Chr X: 194944..195980
X-3	Chr X: 223616..224744
X-4	Chr X: 236336..237310
XI-1	Chr XI: 67491..68573
XI-2	Chr XI: 91575..92913
XI-3	Chr XI: 93378..94567
XI-5	Chr XI: 11779..118967
XII-1	Chr XII: 795787..796720
XII-2	Chr XII: 808805..809939
XII-4	Chr XII:830227..831248
XII-5	Chr XII: 839226..840357

Table S5. Sequences of *E. faecalis* PDH complex genes

Gene	Sequence (5'->3')
pdhA	atggcaaaagcaaagaagcaaaaacctatagatttcaaagaattgatggctaaagtagacgcagatttccctacattccaaatattagatcaagacggtaaaatagttaacgaag atttggtcccagatttgtctgacgaagaattggttgaattgatgacaagaatggtatggtctagagttttggatcaaagatcaaccgcattgaacagacaaggtagattaggtttct ttgctccaaccgcaggtcaagaagcatcccaattagccagtcaattcgctatggaaaaggaagattatttggttgccaggttacagagacgtccctcaattggtagtcaacatggttt gcctttaagagaagctttcttatggtcaagaggtcacgttgctggtaattattacgcagaagatttgaacgccttacccctcaaattataatcggtgcccaatatattcaagctgc aggtgtcgctttgggtttgaagaaaagaggtaaagaaaacgttgtctttacttacacaggtgacggtggttcttcacagggtgacttctacgaagcaattaatttcgccggtgctt accaagctaacggtgttttttattatccaaaacaacggtttcgctatatccacaccaagagaaaaacaaacagccgctaagaccttggcccaaaaagctgttgcagccggtatac caggtatccaagtagatggtatggacccttttagctgtttatgcaatagccaaggaagcaagagattggagtgctgctggtaatggtccagtcttgatcgaaacctttaacttatataga tacggtccacatactttgtccggtgacgaccctacaagatatcgtagtaaggaaatggatgacgaatgggttcaaaaagatcctttaaccagattcagaaagtacttgactgata aaggtttatggtccgaagcaaaggaagaagaaatcatcgaaaagactaaggaagaaattaaggttgctatagcagaagccgataaagctccaaaacaaaaggtctctgacttt ttgaagaacatgttcgaagtacaacctcaaacaatcaaagaacaaatagcattctacgaagccaaagaatcaaagtaa
pdhB	atggcacaaaagactatgatccaagctataactgacgccttagccttagaattagaaaaagatgaaaatgtattgatattcggtgaagatgtaggtaataacggtggtgtttttag agccacagaaggtttgcaagaaaagtttggtgaagatagagttttcgacaccccattggctgaatccggtatagggtggttttagctttttggtttggcattacaaggttatagaccag ttcctgaaatccaatttttcggttttgtcttcgaagtattcgatgaaattgtcggtcaaatggctagaactagatacagaatgggtggtaccagaaatatgccaattactgtaagag caccttttggtggtggtggttcatacaccagaattacactctgataacttggaaggtttaatagcacaatccccctggtgttagagttgtcatcccaagtaatccttatgatgccaagg gtttgttgatctcttcaatcagatccaacgacccagtagtttatttggaacatatgaagttgtacagaagtttcagagaagaagtcccagatgaagcatatgaagtaccttttggaca aagctgcagttactagagaaggtacagatgtctctattataacctacggtgccatggttagagaagctattaaagccgctgattctttagcaaaggacaatatatcagccgaaat cattgatttgagaacagttgctcctttagacgtcgaaaccataatcaactcagttgaaaagactggtagagtcgtagttgtccaagaagcccaaaagcaagctggtgtttggtgca atggtagtttccgaaattagtgaaagagctgtcttgtctttagaagcaccaatagggtagagtatcagcaccagatactatctttccttttcggtcaagccgaaaatatttggttgcct aacgctaaagacatcgaagcaaaggccagagaaattgttgaattttaa
lpd	atggttgttggtgactttgctattgaattggatacagtagttattggtgccggtcctggtggttatgttgccgccatcagagccgcagaaatgggtcaaaaggttgctatcatcga aagagaatacatcggtggtgtttgtttgaacgtcggttgcataccatcaaaagctttgatcgctgcaggtcatcactaccaagaagcacaagattcttcaactttggtgtcacag caaagggtgtaaaattagacttcgccaagacccaagattggaaagacaatacccgttgtcaagtctttgacttcaggtgtaggcatgttgttgaagaaacataaggttgaaatcat cgaaggtgaagccttttcgtagatgaaaacacattgagagttattcacccagactctgcacaaacttactcttttaataacgcaatcgttgccacaggttctagaccaatcgaaa ttcctggttttaagttcggtggtagagttttggattcaactggtggtttgaatttgaaggaagttcctaaaaagttcgtcataatcggtggtggtgttattggtgctgaattaggtggt gcttatgcaaacttgggttctgaagtcacaatcttagaaggttccccaagtattttgcctacctacgaaaaggatatggttaaagtagttactgatgacttcaaaaagaaaaatgta accattgttacttctgccatggctaaagaagctgttgataatggtgactcagtcactgtaaagtatgaagtcaacggtaaagaagaatctgtagaagctgattacgttatggtcac agtaggtagaagaccaaatacgatgacttgggtttagaacaagcaggtgttgaaattggtgaaagaggtttaataccagttgataaccaaggtagaacaaacgtcaaaaacat cttcgctatcggtgacatcgttcctggtgccgctttggctcataaggcatcctatgaagccaaaatcgcagccgaagctattagtggtaaaaaggttgcagtcgattacaaggcc atgccagcagttgccttcacagaccctgaattagcatccgttggtatgaccgttgctgaagcaaaagaagctggtattgaagcaaagggttacaaatttccattcgctggtaacg gtagagcaatctcattggataagactgaaggttttatgagattggtcactacagtagaagacaacgtcattatagggtgcccaaattgctggtgtaggtgcttccgatatgataagt gaattggccttagctatcgaatccggtatgaatgccgaagatattgctttaacaatacatccacaccctagtttgggtgaaattacaatggatacccgcagaattggccttaggttt gcctatacatatctaa
aceF	atggcctatcaattcaagttgcctgacatcggtgaaggtatcgcagaaggtgaaatcgtaaaatggtttgtaaagcctggtgacacaatcaacgaagatgacacttttgttggaag ttcaaaacgataagtctgtcgaagaaatacccatcacctgtaaccggtactgttaaaaatatcgttgtccctgaaggtacagtcgctaacgtaggtgacgttttgattgaaatagac gcaccaggtcacgaagataatgacgctgcaccagccgctcctgcacaagaacaaactccagcccaacctgcagccgttcctactacagaagctgcaggtggtttctttcaatt caaattgccagatataggtgaaggtatagctgaaggtgaaattgtcaaatggttcgtaaaggcaggtgacactataaatgaagatgactctttgttggaagtacaaaacgataag tccgttgaagaaatcccaagtcctgttacaggtaccgtcaaaaatattgtagttccagaaggtactgttgccaatgttggtgacgttttggtcgaaattgacgcaccaggtcataa ttctgccgctccagcagccgctgcacctgctactgatgcaccaaaagccgaagcttccgcaccagccgctagtacaggtgtcgtagcagccgctgatcctaacaagagagtc ttagcaatgccatcagttagacaatatgccagagaaaagatgtagacattacacaagttactgctactggtaaaggtggtagagttatcaaggccgatattgacgcttttgtttc cggtggtagtcaagcagcccctgctactgaagctgcagccacagaagctgcaccaaaagctgaagccgctgcaccaaaagccgctcctaaggcattcacatccgatttgggt gaaatggaaaccagagaaaagatgacccctactagaaaagccattgctaaggcaatggttaattccaaacatacccgctccacacgttacttttgcatgatgaagtagaagttagt aagttgtgtgggatcacagaaagaagtttaaagacgttgcagccgctaatggtacaaaattgaccttcttaccttatgttgtcaaagctttgacatctaccgttcaaaagttcccaatc ttgaacgcatcaatagatgacgcagcccaagaaatagtttacaagaactacttcaacatcggtattgctactgatacagaccatggtttgtacgttccaaacgtcaaaaatgctaa cacaaagtcaatgtttgccatagctgatgaaatcaacgaaaaagctgcattggctattgagggtaaattaaccgcacaagatatgagagacggtaccataactatctctaatata ggttcagttggtggtggttggttcaccccctgttatcaactacccagaagtcgctatcttaggtgtaggtactattgcccaagaaccagtagttaatgctgatggtgaaattgtcgta ggtagaatgatgaagttgtctttgtcattcgatcatagaatcgttgacggtgcaactgcccaaaaagctatgaacaacattaaaagattattggcagacccagaattattattgatg gaaggttga
lplA	atgagatacgtcatcatgcaatccagagacatcagagaaaatttggctaccgaagattacttgttgaacaccttgtctttcgaagaaccattggtcttgttctacattcaagaacca tgcgttatcttgggtagaaatcaaaacgcctacgaagaaattgatttggcttacgctagagaaaagggtatcgttatcactagaagattgtctggtggtggtgctgtttatgatgat ttgggtaatgtctccttctcattcgttgttcaagaaggtcatcaagctttcggtgatttaaggctttcaccaagccaattattgaagccttgcataagatgggtgctactggtgctg aaatttctggtagaaacgatttgttgatcgacggtaaaaagttctctggtaatgctatgtacaccaaaaagggtaagatgtacactcacggtactttgatgtacgatgttgacttgg ctgaagttcaaagagtttgaccgtttccaagaagaagatcgaatctaagggtactaagtccgttagaggtagagttacaaacttaagaccatacttggacggtaaataccaaca

	attgaccatcgaagaattcagaaacagattattgatggaattattcgatgtcgaatccttgaccgaaatcgccgaaaaagaatacgttttgactaagaccgaccaacaagaaatc agaaagttggttgctgaagtctacggtaatgaagcttggatttttggtgaagctccaaagttcaccatcaagaaagaagaaaagttcaagggtggtatcgttgatgctagattga ctgttgaaaaaggtaaaatcatcgaattgactatctacggtgattacttcgctaagaaagaaaccgctgaaatagttgctgctttgttgggtgttgattaccaatactcttctatttg gcaagctttggctgctttcaacttcgaagattatttcgtcaacatcaccaaagaagaattcgtccacttgttggttgactaa
lplA-2	gtcatcttcgtcccaaacgaaaacaacgatccaagagttaacttggccattgaaacttacttgttgaccgaaatgccattggatgaacctatcttgttgttctacatcaacgaacc atccatcatcatcggtagaaatcaaaacaccatcgaagaaatcaacaaagaatacgttgacgaacacggtatccacgttgttagaagattgtctggtggtggtgctgtttatcat gatcacggtaatttgaacttctcattcatcatgccagatgacggtaactcttttagagatttcgctaaggttacccaacctattattcaagccttgcatgatttgggtgttgaaggtg ctgaattgaagggtagaaatgatttggttatcaacgacatgaagttctccggtaatgctatgtatgctaccaatggtagaatgttcgctcatggtactttgatgttcgattccgatat cgatgaagttgtcaacaccttgaaagtcagaaaggacaagattgaatccaagggtatcaagtccgttagatctagagttaccaacatcaagccattcttgtccgaagataagca agaaatgactaccgaagaattcagacaagaaatcttgttgaagatttttcggtgtcgactccatcgatcaagttaagacttatgaattgaccgatcaagattgggctgccattaaca aaatctccgaacaatatttacagaaactgggactggaactacggtaaatctccagcttttaacttggaaagaagacacagattcccaatcggttccattgaaatgaagatgaatgt tgctgatggtgccattcaagaaatcaagatcttcggtgatttttttcggtttgggtgaaatcaaggatgtcgaagatattttgaccggtgttaagtacgataaggcctctttggaaga agccattgatcaaatcgatgtcaagaagtactttggtaacatcgaaaaagaagatttgttgggtttgatctat

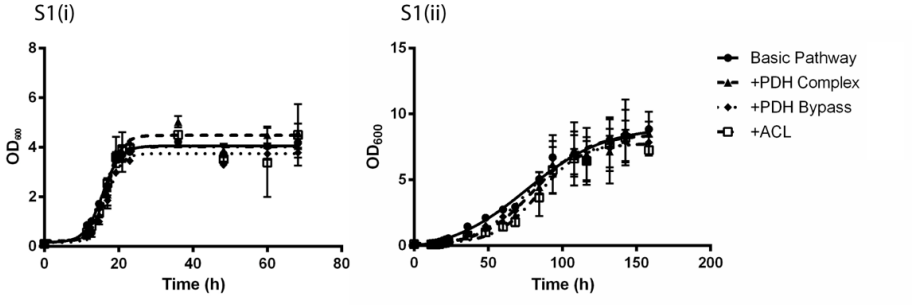


Figure S1. Growth profile of the 3-HP-producing laboratory strains in mineral medium (i), and in simulated fed-batch medium (ii).

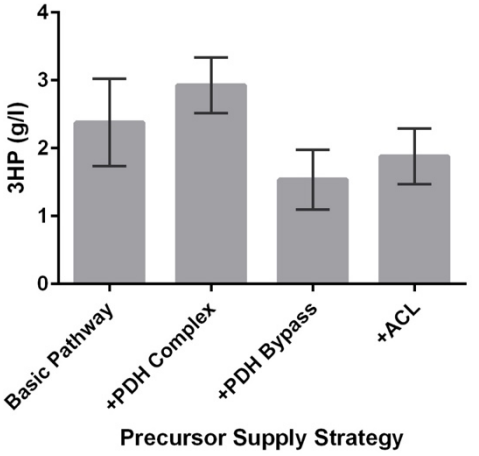


Figure S2. Final titres of 3-hydroxypropionic acid in the engineered industrial Ethanol Red strains.

Supplementary material: The EasyClone-MarkerFree Manual is not included due to length in this thesis but can be found with the online version of this paper.

References

[1] Botstein, D., Chervitz, S. A., Cherry, J. M., Yeast as a model organism. *Science* 1997, *277*, 1259–1260

[2] Borodina, I., Nielsen, J., Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. *Biotechnol. J.* 2014, *9*, 609–620.

[3] Jensen, M. K., Keasling, J. D., Recent applications of synthetic biology tools for yeast metabolic engineering. *FEMS Yeast Res.* 2015, *15*, DOI: 10.1111/1567-1364.12185.

[4] Schiestl, R. H., Petes, T. D., Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 7585–7589.

[5] Jensen, N. B., Strucko, T., Kildegaard, K. R., David, F. et al., EasyClone: Method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2014, *14*, 238–248.

[6] Stovicek, V., Borja, G. M., Forster, J., Borodina, I., EasyClone 2.0: Expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. *J. Ind. Microbiol. Bio- technol.* 2015, *42*, 1519–1531.

[7] Pronk, J. T., Auxotrophic yeast strains in fundamental and applied research. *Appl. Environ. Microbiol.* 2002, *68*, 2095–2100.

[8] Alam, M. T., Zelezniak, A., Mülleder, M., Shliaha, P. et al., The metabolic background is a global player in *Saccharomyces* gene expression epistasis. *Nat. Microbiol.* 2016, *1*, 1–10.

[9] Storici, F., Durham, C. L., Gordenin, D. A., Resnick, M. A., Chromosomal site-specic double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 14994–14999.

[10] Wingler, L. M., Cornish, V. W., Reiterative recombination for the in vivo assembly of libraries of multigene pathways. *Proc. Natl. Acad. Sci. USA* 2011, *108*, 15135–15140.

[11] Kuijpers, N. G., Chroumpi, S., Vos, T., Solis-Escalante, D. et al., One- step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2013, *13*, 769–781.

[12] Kuijpers, N. G., Solis-Escalante, D., Bosman, L., van den Broek, M. et al., A versatile, efficient strategy for

assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb. Cell Fact.* 2013, *12*, 47.

[13] DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X. et al., Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013, 41, 4336–4343.

[14] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012, *337*, 816–821.

[15] Jakociunas, T., Bonde, I., Herrgård, M. J., Harrison S. J. et al., Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* 2015, *28*, 213–222.

[16] Jakociunas, T., Rajkumar, A. S., Zhang, J., Arsovska, D. et al. CasEMBLR: Cas9-facilitated multiloci genomic integration of in vivo assembled DNA parts in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 2015, *4*, 1226–1234.

[17] Stovicek, V., Borodina, I., Forster, J., CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Metab. Eng. Commun.* 2015, *2*, 13–22.

[18] Ronda, C., Maury, J., Jakociunas, T., Jacobsen, S. A. et al., CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 2015, *14*, 97.

[19] Mans, R., van Rossum, H. M., Wijsman, M., Backx, A. et al. CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2015, *15*, DOI: 10.1093/femsyr/fov004.

[20] Laughery, M. F., Hunter, T., Brown, A., Hoopes, J. et al. New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast* 2015, *32*, 711–720.

[21] Horwitz, A. A., Walter, J. M., Schubert, M. G., Kung S. H. et al., Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Systems* 2015, *1*, 88–96.

[22] Jakociunas, T., Jensen, M. K., Keasling, K. D., CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* 2016, *34*, 44–59.

[23] Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E. et al., Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* 2012, *14*, 104–111.

[24] Delneri, D., Tomlin, G. C., Wixon, J. L, Hutter, A. et al., Exploring redundancy in the yeast genome: An improved strategy for use of the cre-*lox*P system. *Gene.* 2000, *252*, 127–135.

[25] Solis-Escalante, D., Kuijpers, N. G. A., van der Linden, F. H., Pronk, J. T. et al., Efficient simultaneous excision of multiple selectable marker cassettes using I-SceI-induced double-strand DNA breaks in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2014, *14*, 741–754.

[26] Nour-Eldin, H. H., Geu-Flores, F., Halkier, B. A., USER cloning and USER fusion: The ideal cloning techniques for small and big laboratories. *Methods Mol. Biol.* 2010, *643*, 185–200.

[27] Gietz, R. D., Schiestl, R. H., High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2007, *2*, 31–34.

[28] Borodina, I., Kildegaard, K. R., Jensen, N. B., Blicher, T. H. et al., Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via beta-alanine. *Metab. Eng.* 2015, *27*, 57–64.

[29] Li, M., Borodina, I., Application of synthetic biology for production of chemicals in yeast *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 2014, DOI: 10.1111/1567-1364.12213.

[30] Kozak, B. U., van Rossum, H. M., Luttik, M. A., Akeroyd, M. et al., Engineering acetyl coenzyme A supply: Functional expression of a bacterial pyruvate dehydrogenase complex in the cytosol of *Saccharomyces cerevisiae*. *MBio* 2014, *5*, e01696–e01614.

[31] Shiba, Y., Paradise, E. M., Kirby, J., Ro, D. K., Keasling, J. D., Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab. Eng.* 2007, *9*, 160–168.

[32] Chen, Y., Bao, J., Il-Kwon, K., Siewers, V., Nielsen, J., Coupled incremental precursor and co-factor supply improves 3-hydroxypropionic acid production in *Saccharomyces cerevisiae*. *Metab. Eng.* 2014, *22*, 104–109.

[33] Jensen, N. B., Borodina, I., Chen, Y., Maury, J. et al., Microbial production of 3-hydroxypropionic acid. WO Patent 198831 A1, 2014.

[34] Kildegaard, K. R., Wang, Z. Chen, Y., Nielsen, J., Borodina, I., Production of 3-hydroxypropionic acid from glucose and xylose by metabolically engineered *Saccharomyces cerevisiae*. *Metab. Eng. Comm.* 2015, *2*, 132–136.

Research Article 2

A comparative study into the effects of pyruvate derived product production pathways on the physiology and metabolism of a Crabtree-negative *Saccharomyces cerevisiae* strain

Mathew M Jessop-Fabre[1], Jonathan Dahlin[1], Vratislav Stovicek[1], Itay Budin[2], Jay D Keasling[1,2,3,4], and Irina Borodina[1]. *Manuscript in preparation* (2018).

[1] The Novo Nordisk Foundation for Biosustainability, Technical University of Denmark, Building 220, 2800 Kongens Lyngby, Denmark
[2] Joint BioEnergy Institute, Emeryville, CA, USA
[3] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[4] Department of Chemical and Biomolecular Engineering & Department of Bioengineering University of California, Berkeley, CA, USA

Abstract

Background

Saccharomyces cerevisiae has long been of industrial importance due to its robustness as a host organism, its tolerance to large-scale fermentations, and because of its high degree of characterisation as a model organism. When fed on high concentrations of glucose, *S. cerevisiae* produces large amounts of ethanol *via* fermentation, even under highly aerobic conditions. To minimise the amount of carbon lost to ethanol, it is possible to produce strains that do not produce any ethanol by deleting pyruvate decarboxylase (PDC) activity. Such Pdc$^-$ mutants accumulate high levels of pyruvate and as such are good potential hosts for the production of pyruvate derived products, such as hydroxy acids.

Results

In this study, we introduce pathways for the production of three common hydroxy acids; lactate, malate, and 3-hydroxypropionic acid (3HP) into an evolved Pdc$^-$ strain (the TAM strain) that is capable of growth on high concentrations of glucose. We observed titres of 16.5 ± 0.1 g L^{-1} pyruvate in the host strain, 30.2 ± 2.2 g L^{-1} lactate in the lactate producing strain, and 26.8 ± 0.3 g L^{-1} malate in the malate producer. Two strains for 3HP production were tested, both sharing the same pathway but with one of the pathway enzymes either using NADH or NADPH as cofactor. The NADH-dependent pathway produced 3.4 ± 0.2 g L^{-1} 3HP, and the NADPH-dependent pathway produced 3.7 ± 0.1 g L^{-1}. Addition of a lactate or malate production pathway into the Pdc$^-$ host strain recovered some growth defects of the TAM strain, most likely due to the replacement of missing NAD^+ regeneration pathways. We employed ^{13}C-based metabolic flux analysis to probe the physiological and metabolic responses of the strains to their engineered pathways. All strains showed considerable flux into the pentose phosphate pathway under shake flask

fermentation conditions – with negligible glucose-6-phosphate isomerase activity, while TCA cycle fluxes are found to remain low across all strains.

Conclusions

An analysis of fermentation characteristics showed that insertion of both the malate and lactate pathways could restore some of the growth performance of the TAM strain, presumably through recycling NADH. This effect was not observed in the 3HP strains, even in the strain capable of regenerating NADH. Through ^{13}C-based metabolic flux analysis, we have shown how the different energetic and redox demands of each pathway cause shifts in flux through central carbon metabolism. We hope that the information gathered here will help in future metabolic engineering efforts in Crabtree-negative strains, and in the production of hydroxy acids in particular.

1 Introduction

Saccharomyces cerevisiae has long been used as a production organism, perhaps most notably for its role in ethanol production. When grown on high concentrations of glucose, yeast will ferment and produce ethanol even in abundant oxygen conditions (Pronk et al., 1996). This is known as the Crabtree effect, and while such a trait has proven useful for bio-ethanol production, ethanol is considered as an unwanted side-product in the bio-based production of many other chemicals (Flikweert et al., 1997; Kim and Hahn, 2015; Nevoigt, 2008). Under fermentative growth on glucose, a large part of the flux from pyruvate passes through the three main pyruvate decarboxylases (PDC) to be converted into acetaldehyde, which is in turn reduced to ethanol *via* alcohol dehydrogenases (ADH), with Adh1p as the major ADH isozyme (de Smidt et al., 2008). Pyruvate decarboxylase activity is conferred through the three major PDC isozymes; Pdc1p, Pdc5p, and Pdc6p (Hohmann, 1991; Schmitt and Zimmermann, 1982; Seeboth et al., 1990).

A successful strategy for the removal of ethanol formation has been to eliminate pyruvate flux into acetaldehyde by deleting all three major PDC isozymes (Flikweert et al., 1996, 1997). Such a Pdc$^-$ strain is hyper-sensitive to glucose and can not grow on high concentrations of glucose without the addition of a C_2 compound such as ethanol or acetate (Flikweert et al., 1999). van Maris et al., (2004a) performed directed evolution on a Pdc$^-$ strain to eliminate its dependency on C_2 compounds, and to improve its glucose tolerance and growth characteristics. While some of the WT growth characteristics were rescued, the resulting strain (referred to as the TAM strain) still has a much slower growth rate than WT; 0.20 h^{-1} *vs.* 0.33 h^{-1} (van Maris et al., 2004a). Oud *et al.*, (2012) later sequenced the genome of the TAM strain and revealed that much of the observed phenotype of the TAM strain was due to a large deletion in the glucose sensing transcriptional regulator *MTH1*. This deletion is hypothesised to reduce the degradation rate of Mth1p, and overexpression of *MTH1* in a non-evolved Pdc$^-$ strain restores growth on high glucose

concentrations and without the need for C_2 supplementation, providing strong evidence for the role of Mth1p as the major driver of the TAM strain's observed phenotype (Oud et al., 2012).

The TAM strain is a pyruvate over-producer, since much of the pyruvate flux is converted to ethanol in Crabtree-positive WT strains, rendering it an attractive host for the production of pyruvate derived chemicals (van Maris et al., 2004a). Pyruvate is a central carbon metabolite, acting as precursor into three major pathways by being converted to; oxaloacetate, mitochondrial acetyl-CoA, and acetaldehyde (Pronk et al., 1996). Pyruvate can serve as precursor for many different heterologous pathways producing compounds of industrial relevance such as the hydroxy acids; malate, lactate, and 3-hydroxypropionic acid (Abbott et al., 2009; Borodina et al., 2015; Ishida et al., 2005; Zelle et al., 2008).

In the present study, we characterised these three different hydroxy acid production pathways and the impact of the added pathways on the metabolism and physiology of the host TAM strain.

Lactate, a three carbon α-hydroxy acid, can be produced from pyruvate in a single enzymatic step by lactate dehydrogenase (LDH) with NADH as the cofactor. This enzyme, most commonly taken from *Lactobacillus* origin, has been shown to produce high levels of lactate in both Pdc^+ and Pdc^- *S. cerevisiae* strains (Branduardi et al., 2006; Colombié et al., 2003; Ishida et al., 2005; van Maris et al., 2004b; Nagamori et al., 2013). The TAM strain has a redox imbalance due to NAD^+ not being regenerated through the formation of ethanol. The addition of the *Lactobacillus plantarum ldh1* gene in a Pdc^- strain can provide the cell with NAD^+, re-balancing the redox state of the cell. However, this redox balancing has been shown to not be able to recover anaerobic growth, with lactate production being highly dependent on oxygen availability (van Maris et al., 2004b). Lactate transport, unlike

ethanol, is assumed to be dependent on ATP, and so cannot be exported under anaerobic conditions where ATP production is limited (van Maris et al., 2004).

High levels of malate production, a four carbon α-hydroxy acid, can be achieved with the overexpression of three genes. The overexpression of the native NADH-dependent malate dehydrogenase *MDH3* lacking the peroxisomal targeting sequence (McAlister-Henn et al., 1995), the native pyruvate carboxylase gene *PYC1*, and the malate transporter from *Schizosaccharomyces pombe MAE1*, allows for high malate titres when expressed in the TAM strain, with reported titres of up to 59 g L^{-1} (Zelle et al., 2008). The activity of Mdh3p in the cytosol also regenerates NAD^+, improving the growth rate compared to the parental TAM strain. Malate transport is assumed to be ATP-dependent, and with the carboxylation of pyruvate to oxaloacetate also requiring ATP, this pathway is more energetically demanding than the lactate production pathway (Abbott et al., 2009; Camarasa et al., 2001; Casal et al., 2008). The export energetics are also different for monocarboxylic and dicarboxylic acids, and export of malate may have a greater ATP requirement that export of lactate (Abbott et al., 2009).

A three carbon β-hydroxy acid, 3-hydroxypropionic acid (3HP), can be produced from at least four different intermediate compounds, with a pathway *via* β-alanine identified as having a high potential for 3HP production (Kumar et al., 2013). This pathway was implemented in a Pdc^+ *S. cerevisiae* strain, proving capable of titres close to 14 g L^{-1} in controlled bioreactor fermentations (Borodina et al., 2015). This pathway requires five enzymatic steps to produce 3HP from pyruvate, three of which require heterologous genes (Figure 1). The final enzymatic step in this pathway, the conversion of malonic semialdehyde to 3HP, can be performed either by a 3-hydroxyisobutyrate dehydrogenase (HIBADH) that uses NADH as cofactor or a 3-hydroxypropionate dehydrogenase (HPDH) that uses NADPH. In Pdc^+ *S. cerevisiae*, the HPDH class of enzymes tends to outperform the HIBADH enzymes, possibly due to the cytosolic ratios of the redox

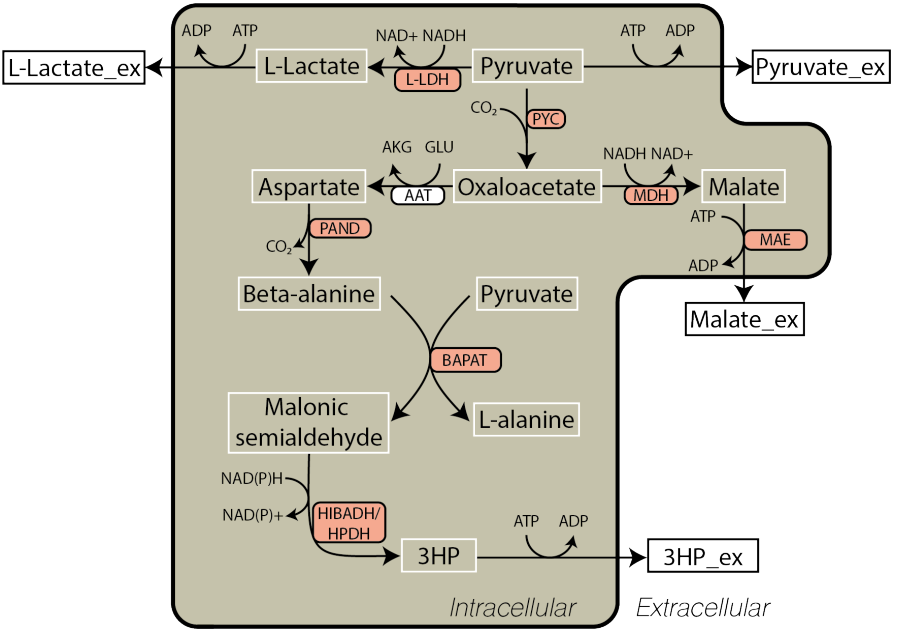


Figure 1. Metabolic routes for the production of three different hydroxy acids in S. cerevisiae. Lactate, malate, and 3HP are produced from pyruvate in the Pdc⁻, pyruvate overproducing host strain (TAM strain) through the implementation of heterologous pathways. Metabolites are shown in white square boxes, while overexpressed genes shown in the oval boxes. Overexpressed enzymes are shown in orange oval boxes. Enzyme abbreviations; L-LDH – L-lactate dehydrogenase, PYC – pyruvate carboxylase, MDH – malate dehydrogenase, MAE – malate transporter, AAT – aspartate amino transferase, PAND – aspartate decarboxylase, BAPAT – β-alanine pyruvate amino transferase, HIBADH – 3-hydroxyisobutarate dehydrogenase (NADH-dependent), HPDH – 3-hydroxypropionic acid dehydrogenase (NADPH-dependent. Metabolite abbreviations; 3HP – 3-hydroxypropionic acid.

cofactors in the cytosol favouring NADPH utilisation (Borodina et al., 2015). The cytosolic ratio of NADH/NAD⁺ in WT *S. cerevisiae* has been found to be very low at ~1/100 (Canelas et al., 2008). However, a Pdc⁻ strain is unable to regenerate cytosolic NAD⁺ through ethanol production, and as such is likely to have a relative excess of NADH, potentially making it more suited than the Pdc⁺ strains for expression pathways using NADH as cofactors, such as lactate, malate, and the HIBADH route of 3HP biosynthesis.

In this study we characterise and compare both the HIBADH and HPDH based pathways for 3HP biosynthesis.

2 Materials and methods

2.1 Strains

Escherichia coli DH5α was used to clone, propagate, and store the plasmids. A derivative of CEN.PK, the TAM strain was a kind gift of Prof. Jack Pronk (Delft Technical University, Holland). A description of each strain used in this study is shown in Table 1. For the respiration experiments, the CEN.PK Pdc$^+$ strain (MATaURA3 HIS3 LEU2 TRP1 MAL2-8c SUC2) was used as reference. This was a kind gift from Peter Kötter, Johann Wolfgang Goethe University, Frankfurt, Germany.

2.2 Plasmid construction and yeast transformation

Each of the plasmids used to create the hydroxy acid producing strains were constructed using the EasyClone vector set for chromosomal gene integrations with auxotrophic selection markers as described in (Jensen et al., 2014). Yeast promoter biobricks were amplified from the genomic DNA of CEN.PK113-7D *via* PCR, using primer overhangs for cloning as described in (Jensen et al., 2014). ST7031 was built with the insertion of the *Lactobacillus plantarum ldh1* under the control of the *TDH3* promoter for the conversion of pyruvate to lactate. Strain ST5413 was constructed from the TAM strain with the insertion of *MDH3* lacking the SKL mitochondrial targeting sequence, under control of the *TDH3* promoter. This gene was amplified from the genomic DNA of *S. cerevisiae* with primers pr7011 and pr7012 to amplify the region lacking the mitochondrial targeting sequence. The native *S. pombe* malate transporter gene *mae1* was also inserted under the control of the *TEF1* promoter, along with the native *PYC1* and *PYC2* genes under the control of the *TEF1* and *PGK1* promoters, respectively. ST2779 and ST2780 were engineered for 3HP

production, following the designs of (Borodina et al., 2015). Native *PYC1* and *PYC2* were overexpressed (under the control of the *TEF1* and *PGK1* promoters), along with the aspartate-1-dehydrogenase gene (*panD*) from *Tribolium castaneum* under the control of the *TEF1* promoter, and the β-alanine-pyruvate amino transferase (BAPAT) from *Bacillus cereus (yhxA)* under the control of the *TEF1* promoter. For the final enzymatic step of the 3HP biosynthetic pathway in ST2779, a 3-hydroxypropionate dehydrogenase (HPDH) from *E. coli* (*ydfG*) that uses NADPH as cofactor was expressed under the control of the *PGK1* promoter. For ST2780, a 3-hydroxyisobutyrate dehydrogenase (HIBADH) from *Pseudomonas putida* (*hibdh)* was inserted that uses NADH as cofactor, under the control of the *PGK1* promoter. Heterologous genes were codon optimised and synthesised by GeneArt (Thermo Fisher Scientific, USA). Sequences of these genes can be found in the supplementary materials. Details and sequences of the genes used to create the 3HP strains can be found in (Borodina et al., 2015). PCR amplification was performed on each gene, using the primers as described in the supplementary materials with corresponding primer overhangs to be compatible with the promoter biobricks (Jensen et al., 2014).

After successful amplification of the required promoters and genes, they were USER cloned into EasyClone vectors (details of which are found in the supplementary materials) previously digested with SfaAI (Thermo Scientific, USA) and Nb.BsmI (New England Biolabs, USA) (Jensen et al., 2014). USER cloning (New England Biolabs, USA) was performed as previously reported, and cloned plasmids were transformed into the *E. coli* cloning strain DH5α *via* heatshock at 42 ˚C for 1 minute (Jensen et al., 2014; Nour-Eldin et al., 2010). Cells were plated on lysogeny broth (LB) agar plates with 100 μg ml^{-1} ampicillin and incubated at 37 ˚C overnight. The next day, individual colonies were tested by PCR for the correct sized plasmid insertions using the verification primers outlined in the supplementary materials. All plasmids were additionally confirmed by Sanger sequencing to verify the correct sequences. Confirmed plasmids were stored as glycerol stocks kept at -80 ˚C.

Plasmids for transformations were first digested with NotI (Thermo Scientific, USA) to produce linear DNA fragments. The fragments containing the genes to be transformed were purified from the vector backbone by gel electrophoresis followed by band excision and gel purification. All yeast transformations were carried out using the lithium acetate chemical transformation protocol as described in (Gietz and Schiestl, 2007). After heatshock, the cells were plated directly onto synthetic complete (SC) agar plates lacking the relevant amino acid required for selection. After 2-3 days, plates were examined for growth, and correct isolates were identified *via* performing PCR on individual colonies using the verification primers outlined in (Jensen et al., 2014).

Complete lists of plasmids, biobricks, and primers used in this study are included in the supplementary materials.

2.3 Media

For selection of *E. coli*, LB agar plates were made to contain 100 μg mL^{-1} ampicillin. For plasmid propagation, LB broth was used with the same concentration of ampicillin.

To select for *S. cerevisiae* transformants and for solid culture, SC agar plates were prepared with pre-mixed amino acid drop-out powders, with a final glucose concentration of 20 g L^{-1}. For the creation of pre-cultures, liquid SC media was prepared with pre-mixed amino acid drop-out powders, also with a final glucose concentration of 20 g L^{-1}. For the respiration experiments, liquid complete synthetic mixture (CSM) was prepared from drop-out powders purchased from Sunrise Science (USA), with 20 g L^{-1} glucose.

S. cerevisiae shake flask fermentations were performed with batch production media optimised for acid production, prepared according to (Zelle et al., 2008), with 100 g L^{-1} glucose as the carbon source.

2.4 Cultivation conditions

For shake flask fermentation experiments, single colonies of each strain were taken from plates and inoculated into 5 mL of liquid SC media with appropriate amino acid selection in 12 mL culture tubes and incubated at 30 °C, 250 rpm for 2 days. Shake flasks were prepared as follows; 5 g $CaCO_3$ was added to each 500 mL baffled shake flask before the flasks were autoclaved. Once cool, 100 mL of the sterile batch production media was added to each shake flask. The prepared flasks were inoculated from the precultures to a starting OD_{600} of 0.5. Fermentations were carried out at 30 °C, and shaking set to 250 rpm, for 5 days with samples taken every ~12 hrs and measured for optical density at 600 nm in an Implen P300 spectrophotometer (dilutions were made to keep measurements within the linear range of the equipment). Part of each sample was centrifuged at 11,000 xg for 5 min, and the supernatant was stored at -20 °C until HPLC analysis. Each strain was fermented in triplicate in a shaking incubator for 120 hrs at 30 °C and 250 rpm.

2.5 ^{13}C-based metabolic flux analysis

For ^{13}C-based metabolic flux analysis (MFA), the same media was used as in the shake flask fermentations, except that ammonia was used as the nitrogen source instead of urea, no $CaCO_3$ was added (to prevent non-labelled carbon interference), and ^{13}C-labelled glucose was used to a final concentration of 10 g L^{-1}. 250 mL shake flasks were filled with 10 mL of media, and inoculated with the strains to a starting OD_{600} of 0.06. Samples for analysis were taken once the OD_{600} of the cells had reached 2.00. 0.3 mg CDW was removed and washed in 0.9% NaCl before storage at -80 °C. Each strain was cultured in

quadruplicate, with one culture grown on 20% U-^{13}C and 80% unlabelled glucose, and the three other cultures grown on 20% U-^{13}C and 80% 1-^{13}C glucose. Labelled glucose was purchased from Euriso-Top (France). The same cultivation was performed in triplicate without labelled glucose, where samples were taken regularly during growth and measured by HPLC for the rates of secreted products, used to constrain the stoichiometric model for metabolic flux analysis.

The frozen cell pellets were resuspended in 150 µl HCL (6M) before being transferred to a silanized glass vial. The samples were left to hydrolyse at 105 °C for 6 hrs, before the remaining liquid was evaporated at 80 °C. The pellets were resuspended in 30 µl acetonitrile. Derivatisation was performed by adding N-methyl-N-*tert*-butyldimethylsilyl-trifluoroacetamide to a final ratio of 1:1, and incubating at 85 °C for 1 hr. GC-MS analysis of the amino acid content was performed on the derivatised samples as previously reported (Kildegaard et al., 2016).

The raw GC-MS data were corrected before flux determination was carried out by the iMS2Flux software for the PERL programming language (Poskar et al., 2012). Steady-state metabolic flux analysis was performed with the INCA toolbox for MATLAB (Young, 2014). A previously reported model metabolic network for *S. cerevisiae* was adjusted to fit the central carbon metabolism of each acid producing strain and used as the model network input for INCA (Wasylenko and Stephanopoulos, 2015). The full list of included reactions is included in the supplementary information.

The calculated fluxes from INCA were used to constrain a modified version of the iMM904 genome scale metabolic model (GEM) of *S. cerevisiae* (Mo et al., 2009). This model was modified by deleting the PDC reaction to eliminate ethanol production, and the heterologous pathways for each strain were added to the reactions in the model, producing a different model for each strain. These models were used to determine the remainder of

the fluxes using parsimonious flux balance analysis (pFBA) implemented in the Cameo toolbox for Python (Cardoso et al., 2018; Lewis et al., 2010).

2.6 Analytical methods

The concentrations of extracellular metabolites were measured by HPLC. The supernatant was removed and transferred to Nunc 96-well plates with rubber sealing top (Thermo Scientific, USA). 30 μL of each sample was analysed for 30 minutes with an Aminex HPX-87H ion exclusion column (Bio-Rad, USA) at 60 °C, with a flow rate of 600 μL min^{-1}. 5 mM H_2SO_4 was used as eluent, and the compounds were detected with a Dionex RI-101 Refractive Index Detector at 45 °C. UV was detected with a DAD-3000 Diode Array Detector at 210 nm (Dionex, USA). For analysis of 3HP, the same protocol was run apart from that 1 mM H_2SO_4 was used as eluent, with a longer run time of 45 minutes. Due to the overlapping spectra of pyruvate and malate in this method, malate concentrations were verified using the L-malic acid assay kit (K-LMAL-116A, Megazyme, Republic of Ireland). Pyruvate was verified using a pyruvic acid assay kit (K-PYRUV, Megazyme, Republic of Ireland).

2.7 Respiration monitoring

To analyse the respiration characteristics of the hydroxy acid production strains, each strain was tested in a respiration activity monitoring system (RAMOS) (Anderlei et al., 2004). Individual colonies of each strain were inoculated into 5 mL of CSM lacking the appropriate amino acids. The following day the RAMOS was calibrated and checked for correct function following the manufacturer guidelines. The RAMOS shake flasks were inoculated to a starting OD_{600} of 0.2 and set to shake at 300 rpm with the temperature set to remain constant at 30 °C. The fermentations lasted for 48 hrs, with oxygen and carbon

dioxide automatically measured every 30 mins. Samples were taken for analysis and tested for optical density every ~12 hrs. Each experiment was performed in triplicate.

2.8 Statistics

The fermentation data in this study were analysed, and statistics calculated using GraphPad Prism software version 7 (GraphPad Software, Inc.). Significance was calculated with Tukey's multiple comparisons test for one-way ANOVA. All values are expressed as the mean average ± SD.

2.9 Chemicals

Unless otherwise stated, all reagents and chemicals were purchased from Sigma-Aldrich, USA.

3 Results and discussion

3.1 Physiological characterization of hydroxy acid-producing Pdc$^-$ *S. cerevisiae*

We constructed different hydroxy acid producing strains, with an evolved glucose tolerant Pdc$^-$ strain as the host – referred to as the TAM strain (van Maris et al., 2004a). The un-engineered host produces and excretes high levels of pyruvate, and so we chose three hydroxy acids that are commonly produced *via* pyruvate; lactate, malate, and 3HP. We reconstructed previously published designs for high level production of lactate, and malate from the TAM strain (Lui and Lievense, 2005; Zelle et al., 2008). Two different versions of the β-alanine route to 3HP biosynthesis were constructed, each utilising a different cofactor in the final enzymatic step of the pathway (Borodina et al., 2015), so that five total

Strain	ST2803	ST7031	ST5413	ST2779	ST2780
Description	Parental TAM strain	Lactate producer	Malate Producer	3HP producer (HPDH)	3HP producer (HIBADH)
Max Product Titre (g L^{-1})	16.5 ± 0.1	30.2 ± 2.2	26.8 ± 0.3	3.7 ± 0.1	3.4 ± 0.2
Product Yield (mol mol^{-1} glucose)	0.340 ± 0.004	0.600 ± 0.091	0.363 ± 0.010	0.070 ± 0.005	0.067 ± 0.007
Glucose Uptake Rate (mmol gDW^{-1} h^{-1})	0.640 ± 0.125	2.317 ± 0.696	4.145 ± 0.519	0.573 ± 0.080	0.497 ± 0.063
Final OD_{600}	63.9 ± 2.2	86.2 ± 2.4	58.6 ± 2.2	51.6 ± 0.5	56.5 ± 1.6
Maximum specific growth rate (h^{-1})	0.080 ± 0.006	0.131 ± 0.007	0.096 ± 0.007	0.063 ± 0.003	0.059 ± 0.004
Succinate (g L^{-1})	0.17 ± 0.06	0.42 ± 0.02	3.97 ± 0.08	0.70 ± 0.06	0.91 ± 0.09
Glycerol (g L^{-1})	6.00 ± 0.05	0.25 ± 0.01	2.54 ± 0.06	2.07 ± 0.20	2.93 ± 0.18

Table 1. Fermentation parameters of the hydroxy acid production strains. Results in the table show the analysed parameters of 120 hour shake flask fermentations. Batch mineral media was used, with a starting glucose concentration of 100 g L^{-1}. Results are from fermentation run in triplicate ± SD. Max product titre refers to the main or target secreted metabolite of each strain; ST2308 – Pyruvate, ST7031 – Lactate, ST5413 – Malate, ST2779 and ST2780 – 3HP.

strains were tested against each other (see Figure 1 for details on the implemented pathways).

To first characterise the production performances of the hydroxy acid producing strains, we performed shake flask fermentations using a batch media composition that has previously been shown to be suitable for the production of malate to high titres in the TAM strain (Zelle et al., 2008). These were performed with 100 g L^{-1} of glucose as the carbon source. The results of these fermentations are shown in Figure 2, with Table 1 listing the key fermentation parameters of each strain.

Biomass accumulation was similar across most of the strains, apart from the lactate producing strain (ST7031) that showed a significantly higher final optical density than all

other strains ($p < 0.001$). The NADPH-dependent 3HP producer (ST2779) showed a slightly lower biomass accumulation ($p < 0.05$) compared to the parental strain (ST2803). Growth rates, however, showed considerable variation with ST7031 producing the fastest growth rate of 0.131 ± 0.007 h^{-1}. Both ST7031 and the malate production strain (ST5413 - growth rate of 0.096 ± 0.007 h^{-1}) showed significantly improved growth rates over ST2803 (0.080 ± 0.006 h^{-1}), while both ST2779 and ST2780 showing significantly lower growth rates than ST2803 (0.063 ± 0.003, and 0.059 ± 0.004 h^{-1} for the ST2779 and ST2780 strains, respectively). The insertion of NAD^+ regeneration reactions improved the growth rates in both ST7031 and ST5413. The extra ATP requirement of malate over lactate production may be the main factor contributing to the lower growth rate of ST5413 compared to ST7031. There was no significant difference between the growth rates of the 3HP producing strains. A slight growth improvement of the HIBADH (ST2780) strain over the HPDH strain (ST2779) may have been expected due to ST2780 regenerating NAD^+ during 3HP synthesis, but no such difference between the strains was observed. The lower than ST2803 growth rates suggest that both 3HP pathways are deleterious to the fitness of the host, and this effect together with the much lower titres of 3HP compared to the other hydroxy acids, may cancel any positive effects provided by NAD^+ regeneration. Glucose was consumed rapidly in both ST7031 and ST5413, but with ST5413 having consumed all of the glucose ~20 hours earlier than ST7031. Both 3HP producing strains, and the parental strain took much longer to deplete the glucose in the media, occurring only at the end of the fermentations.

ST2803, ST7031, and ST5413 saw growth coupled production of their acids, which stopped on glucose depletion. ST7031 and ST5413 both produced high levels of their products, with a final measured lactate and malate concentrations of 30.2 ± 2.2, and 26.8 ± 0.3 g L^{-1} respectively. This was higher than the pyruvate production in the parental strain, which had a maximum titre of only 16.5 ± 0.1 g L^{-1}.

Pyruvate yields on glucose were reported in the original TAM publication to be 0.92 mol mol^{-1} glucose (van Maris et al., 2004a). In this study we recorded yields of only 0.34 mol mol^{-1} glucose. This difference may be due to the original paper using a multiple glucose feeding approach to maximise pyruvate yields. Lactate production from a TAM strain expressing a bacterial LDH has been reported to produce up to 50 g L^{-1} lactate in shake flask fermentations with only 70 g L^{-1} glucose, a yield of 1.44 mol mol^{-1} glucose (Lui and Lievense, 2005). This is much higher than the yields we report (0.60 mol mol^{-1} glucose), although there are multiple experimental differences, perhaps the most important being that in our study we use a chromosomally integrated LDH gene present in a single copy, whereas in the previous work they express an LDH from a high copy number plasmid, as well as this study using different fermentation parameters (Lui and Lievense, 2005). Although the construction of ST5413, and the media used for fermentation, were almost identical to that previously published by (Zelle et al., 2008), we achieved a slightly lower yield of malate (0.36 *vs.* 0.42 mol mol^{-1} glucose). This difference may be due to differences in glucose or $CaCO_3$ concentrations used between the studies, as well as the difference in expression methods between the two studies (single gene integration into the yeast chromosome in the present study *vs.* high copy number plasmid expression in the original publication). These figures are also from 300 hour long fermentations, whereas this study ran 120 hour fermentations (by which point the glucose appeared to have been fully consumed). Pdc^+ *S. cerevisiae* expressing the HPDH version of the 3HP pathway produced a yield of 0.09 mol mol^{-1} glucose in batch media conditions, slightly higher than the same pathway expressed in the TAM strain – ST2779 (0.07 mol mol^{-1} glucose) (Borodina et al., 2015). The poorer performance of the TAM based strains compared to the Pdc^+ host is surprising, and perhaps due to the media conditions not being optimal for 3HP production. However, the 3HP titres were much higher than those previously published under batch fermentation conditions (Borodina et al., 2015).

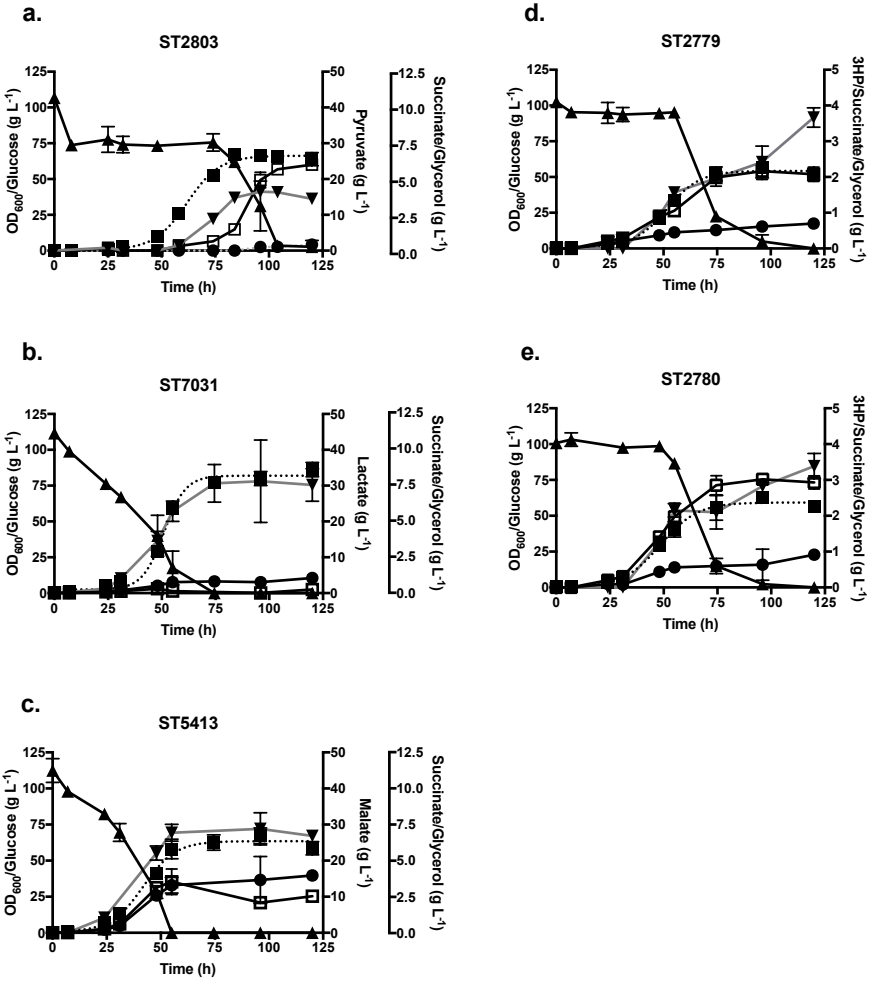


Figure 2. Fermentation profiles of the hydroxy acid production strains. Time-course fermentations showing a) ST2803 - parental TAM strain, b) ST7031 - TAM strain with lactate production pathway, c) ST5413 - TAM strain with malate production pathway, d) ST2779 - TAM strain with HPDH-dependent 3HP pathway, e) ST2780 - TAM strain with HIBADH-dependent 3HP pathway. Each fermentation was run for 120 hrs and in triplicate. Values shown are the mean average for each time point ± SD. Glucose (▲)*, Product* (▼)*, Glycerol* (□)*, Succinate* (●)*, OD_{600}* (■)

ST2779 and ST2780 showed continued 3HP production into stationary phase. The two 3HP strains produced similar levels of 3HP by the end of fermentation, 3.7 ± 0.13, and 3.4 ± 0.17 g L^{-1} respectively. When these enzymes were expressed in a Pdc^+ strain, there was a greater difference between the performance of these two enzymes, with the strain carrying HPDH producing ~4-fold higher titres of 3HP than a strain carrying the HIBADH enzyme (Borodina et al., 2015). This observation is consistent with expectations as the Pdc^- strain exhibits an NADH imbalance, and therefore the NAD^+ regenerating pathway in ST2780 is expected to perform relatively better in the TAM strain.

ST2803 produced barely detectable levels of secreted by-products, with the exception of glycerol. Glycerol production occurred late in the fermentation, with ~6 g L^{-1} measured at the end of 120 hours. Glycerol is formed from the splitting of fructose-1,6-bisphosphate into dihydroxy-acetone-phosphate (DHAP), and glyceraldehye-3-phosphate. DHAP can be converted to glycerol, yielding 1 mol NAD^+, and glyceraldehyde-3-phosphate can be converted to pyruvate, yielding 1 mol NADH and 2 mol ATP. Over accumulation of pyruvate may then shift some flux into glycerol production. It may also be a consequence of redox imbalance in the cell, regenerating more NAD^+. Glycerol production is lower in most of the other strains, apart from ST5413, which produced a maximum of ~4.5 g L^{-1}, which then partly appears to have been re-assimilated. The final molar yield of lactate in ST7031 was ~66% greater than that of malate in ST5314 (0.60 *vs.* 0.36 mol mol^{-1} glucose), potentially leading to higher level of available NAD^+ in the ST7031. Unlike lactate, malate can feed in to the TCA cycle, and also into the glyoxylate shunt (Zelle et al., 2008). Overproduction of malate may cause perturbations, where by-products accumulate but are not able to be processed due to low NAD^+ levels. This may account for the higher secreted succinate production, produced in both the TCA cycle and glyoxylate shunt, which is observed in only in ST5413 (final concentration of 4 g L^{-1}).

Differences in by-product secretion are found between the two 3HP producing strains. Both produced similar levels of succinate (~0.8 g L^{-1}), but ST2780 produces 50% more glycerol than ST2779 (3 g L^{-1} and 2 g L^{-1} respectively). The cause for this is not readily apparent, as the conversion of DHAP to glycerol-3-phosphate generates NAD^+, as does the final 3HP catalysing step of the NADH-dependent producer. It may then be expected that the NADPH dependent strain would produce more glycerol to compensate for a lack of NAD^+ generation through ethanol production.

3.2 RAMOS

The respiration activity monitoring system (RAMOS) is capable of measuring both the oxygen and carbon transfer rates (OTR and CTR) during microbial fermentations, and from this determine the respiratory quotients (RQ) of the strains (Anderlei et al., 2004). Figure S1f shows the fermentation profile of the Pdc^+ CEN.PK strain, where it is apparent that the OTR and CTR are highly uncoupled, leading to a high RQ during glucose assimilation. ST2803 shows a tightly linked OTR and CTR as it progresses through the fermentations, with an RQ of 1. All of the tested strains had near identical OTR's during the first 10 hours of the fermentations. ST7031 was the first Pdc^- strain to show a slow down of OTR, beginning to slow at ~18 hours (Figure 3). However, later in the fermentation, the OTR of ST7031 begins to rise again. At between 28.5 and 31 hours, the RQ is 0.79 ± 0.04, which could suggest that it is assimilating a secreted product such as glycerol (with an RQ of 0.86) along with glucose. There is also a second similarly sized increase in activity at the end of the fermentations (Figure 3). The marked difference in the ST7031 OTR profiles as compared to ST5413 is interesting, as both produce additional NAD^+. Additionally, ST7031, similar to the Pdc^+ strain, is able to metabolise all of the glucose with lower levels of total oxygen consumption, while ST5413 retains the higher demand for oxygen that is found in the parental TAM strain (Figure S3).

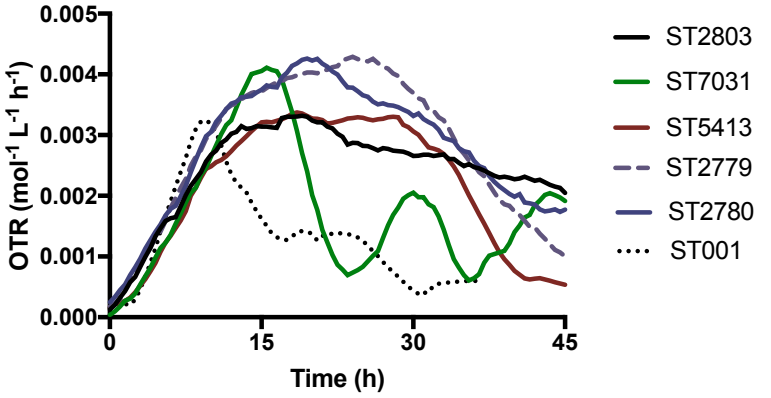


Figure 3. Respiration characteristics of the hydroxy acid production strains. Average oxygen transfer rate (OTR) profiles over the course of the fermentations were measured for each production strain, the parental TAM strain, and a Pdc$^+$ WT CEN.PK strain (ST001) for reference. For strain references, refer to Table 1. Results shown are the mean average of fermentations run in triplicate.

3.3 ^{13}C-based metabolic flux analysis

To further characterise hydroxy acid production in Pdc$^-$ yeast, we employed ^{13}C-based metabolic flux analysis combined with genome scale modelling to identify flux distribution variation between hydroxy acid producing strains. Each strain was cultivated in defined liquid batch media that has previously shown to be beneficial for malate production, containing ^{13}C-labelled glucose as the carbon source (Zelle et al., 2008). The strains were cultivated in shake flasks, and samples taken when each strain reached mid-exponential phase. Once collected, samples were analysed by GC-MS and the fluxes were calculated using the iMS2flux, and INCA programs (Poskar et al., 2012; Young, 2014). The iMM904 genome scale model of *S. cerevisiae* was constrained with the calculated fluxes to estimate flux distribution for each strain on a global metabolic level (Figure 4). Glucose uptake rates for ST2803 and both 3HP producing strains were comparable, but ST7031 and ST5413 showed 3.5 fold and 6.5 fold faster uptake rates, respectively (Table 1). Growth rates were likewise higher on both ST7031 and ST5413, showing that both pathways help to

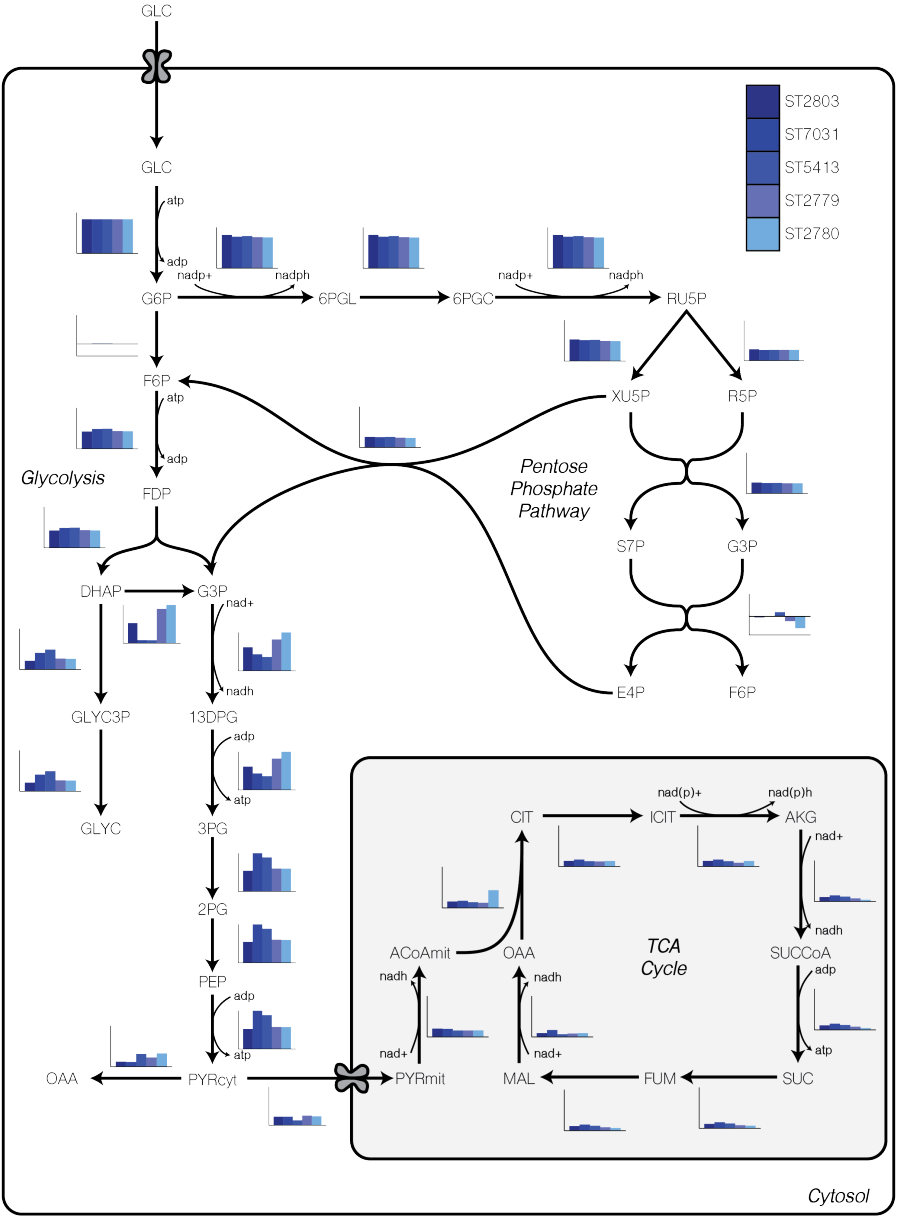

GLC
GLC
ST2803
ST7031
ST5413
ST2779
ST2780
atp
adp
nadp+
nadph
G6P
6PGL
6PGC
RU5P
F6P
XU5P
R5P
Glycolysis
Pentose
Phosphate
Pathway
FDP
S7P
G3P
DHAP
G3P
nad+
nadh
E4P
F6P
GLYC3P
13DPG
GLYC
3PG
2PG
PEP
OAA
PYRcyt
CIT
ICIT
AKG
nad(p)+
nad(p)h
ACoAmit
OAA
TCA
Cycle
SUCCoA
PYRmit
MAL
FUM
SUC
Cytosol

Figure 4 (Over page). ^{13}C-based metabolic flux analysis of the hydroxy acid production strains. Crabtree negative S. cerevisiae strains (TAM – ST2803) producing different hydroxyacids; lactate (ST7031), malate (ST5413), and 3HP (ST2779 and ST2780) were analysed for their relative flux distributions. Bar charts next to each reaction show the relative flux values for each strain (normalised over glucose uptake rate). GLC glucose, G6P glucose-6-phosphate, 6PGL D-6-phospho-glucono-δ-lactone, 6PGC 6-phospho-D-gluconate, RU5P ribulose-5-phosphate, XU5P xylulose-5-phosphate, R5P ribose-5-phosphate, S7P sedoheptulose-7-phosphate, G3P glyceraldehyde-3-phosphate, E4P erythrose-4-phosphate, F6P fructose-6-phosphate, FDP fructose-1,6-diphosphate, DHAP dihydroxy-acetone-phosphate, GLYC3P glycerol-3-phosphate, GLYC glycerol, 13DPG 1,3-diphosphateglycerate, 3PG 3-phosphoglycerate, 2PG 2-phosphoglycerate, PEP phosphoenolpyruvate, PYRcyt pyruvate (cytosolic), ACALD acetaldehyde, EtOH ethanol, AC acetate, PYRmit pyruvate (mitochondrial), ACoAmit acetyl-CoA (mitochondrial), CIT citrate, ICIT isocitrate, AKG α-ketoglutarate, SUCCoA succinyl-CoA, SUC succinate, FUM fumarate, MAL malate, OAA oxaloacetate.

recover some growth capabilities. Under respiro-fermentative growth, as is expected to be experienced during shake flask cultivations, flux through the pentose phosphate (PP) pathway and TCA cycle are low in Pdc$^+$ WT strains (Gombert et al., 2001; Maaheimo et al., 2001). In comparison to Pdc$^+$ WT strains, all strains analysed in the present study diverted most of the carbon flux (>90%) into the PP pathway. However, flux through the TCA cycle remained low, especially in both 3HP producing strains.

Both 3HP strains showed higher relative fluxes into glycerol-3-phosphate from DHAP, and from the reverse reaction of erythrose-4-phosphate and fructose-6-phosphate. The increased flux into glycerol-3-phosphate then appears to be diverted away from the central pathway, and into glutamate biosynthesis from 3-phosphoglycerate. ST7031 and ST5413 are able to keep a high proportion of their carbon flux through 3-phosphoglycerate to pyruvate, presumably due to the presence of a suitable carbon sink. From pyruvate, both the 3HP strains show flux ratios similar to ST5413 for oxaloacetate synthesis, suggesting that *PYC1* and *PYC2* overexpression is correctly functioning to push flux into oxaloacetate in these strains, but they are unable to convert this additional flux into 3HP production. ST2780 shows higher pyruvate carboxylase flux than the ST2779, but both strains show near identical fluxes through the aspartate transaminase step, with the reaction only processing around half of the flux supplied from the pyruvate carboxylase reaction. This

could suggest that *AAT2* overexpression, or work to remove regulation of this gene may improve the flux to aspartate, with positive results reported previously for the overexpression of the native *AAT2* in the Pdc$^+$ background (Borodina et al., 2015). However, it is also clear that there is further room for improvement in the final three enzymatic steps of the pathway, as only a relatively low flux is carried from aspartate (see Figure S4). The impaired growth of both 3HP strains may be due in part to a toxic build up of malonic semialdehyde, with the detoxofication by HPDH/HIBADH unable to proceed quickly enough to allow the cell to grow rapidly (Dalwadi et al., 2017; Kim et al., 2010). Investigations into the transport mechanisms of 3HP from the cytosol could help to pull flux from malonic semialdehyde and restore growth. Another strategy may be to engineer a malonic semialdehyde sensor to dynamically control flux through the pathway, similar to previous work carried out on 3HP biosensors (Hanko et al., 2017; Rogers and Church, 2016).

Transcriptomics can provide valuable insight into how a strain may try to adapt to genetic modifications. A deeper understanding of stresses faced by each strain during fermentation, and solutions for minimising the stress factors, may be provided through transcriptomic analysis. Future work in this area may reveal where regulatory networks are acting to prohibit higher product formation in each strain.

4 Conclusions and future perspectives

In this study, we have provided a characterisation of hydroxy acid production pathways in a Crabtree-negative, pyruvate overproducing *S. cerevisiae* strain – the TAM strain. Introduction of a lactate or malate production pathway restores some of the growth defects found in the TAM strain by re-supplying cytosolic NAD$^+$. The energetic requirements of lactate and malate transport, combined with the mutation found in *MTH1*, prevents full restoration of growth rate to that of Pdc$^+$ *S. cerevisiae*. However, NADH-

dependent production of 3HP in ST2780 was unable to provide any growth restoration, and both 3HP producing strains had reduced growth rates compared to ST2803.

The presence of malate or lactate as a carbon sink allowed these strains to achieve a higher relative flux through lower glycolysis than the TAM strain. This was not the case for the 3HP producers, and so further research must be undertaken to improve the downstream capacity of the 3HP pathways.

Acknowledgement

This work was funded by the Novo Nordisk Foundation. VS and IB acknowledge funding for the BioREFINE-2G project by the European Commission in the 7th Framework Programme (Project no. FP7-613771). JD acknowledges funding from an International Training Network PAcMEN Marie Curie grant. The authors declare no financial or commercial conflict of interest.

Supplementary Information

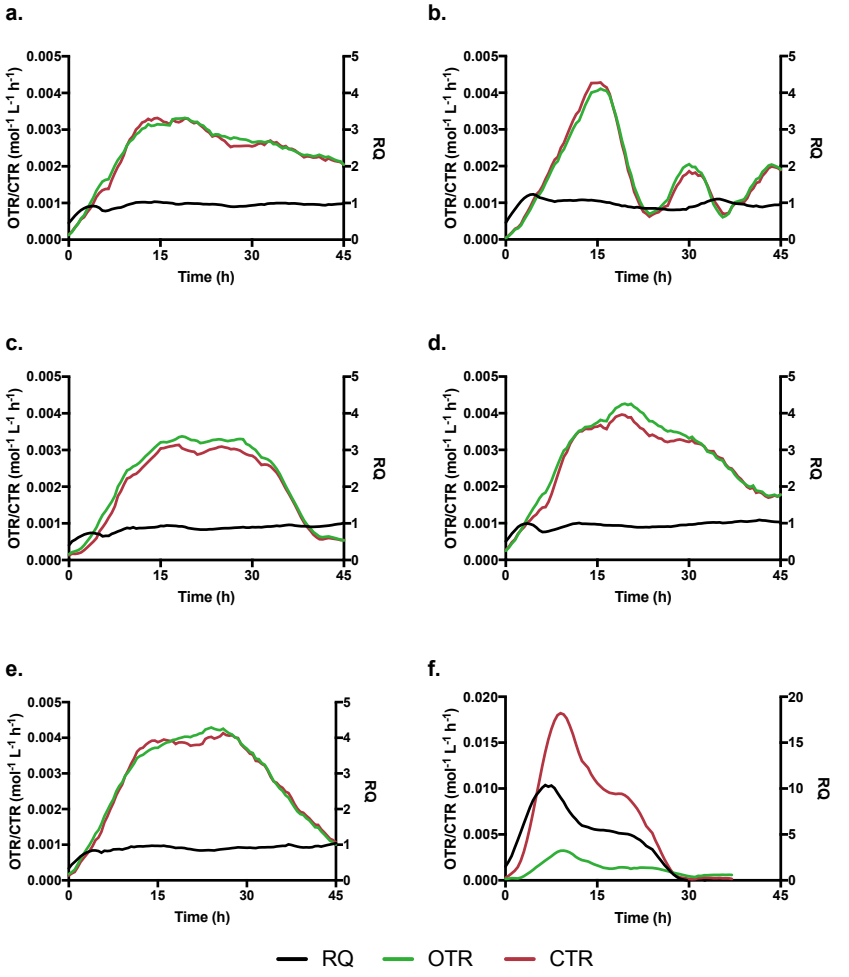


Figure S1. Respiration profiles of each hydroxy acid production strain. a) ST2803 - Parental TAM strain, b) ST7031 - Lactate producing strain, c) ST5413 - Malate producing strain, d) ST2779 - NADPH-dependent 3HP producing strain, e) ST2780 - NADH-dependent 3HP producing strain, f) ST001 - CEN.PK Pdc^{+} WT strain. Respiratory Quotient is designated by the black lines, oxygen transfer rate by the green, and carbon dioxide transfer rate by red. Experiments were carried out in triplicate.

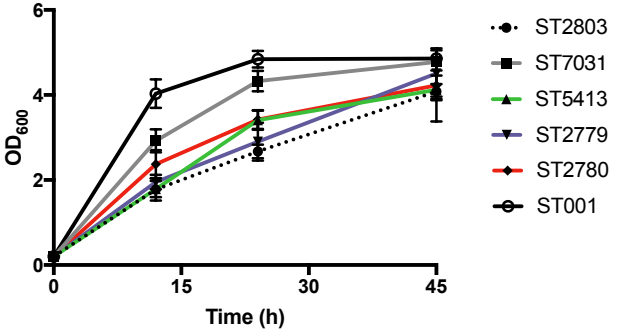


Figure S2. Optical density profiles of the respiration experiments. The measured OD_{600} for each strain was taken from triplicate experiments run in the RAMOS system to measure the respiration characteristics during fermentation. Starting OD_{600} was set to 0.2.

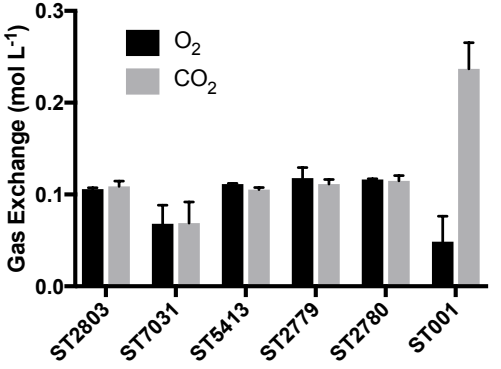


Figure S3. Total gas exchanges after 37.5 hours. The volume of oxygen taken up by the strains in the respiration experiments is shown by the black bars, and the total volume of carbon dioxide released is shown by the grey bars. The wild-type CEN.PK Pdc^+ - ST001, strain is used as a control.

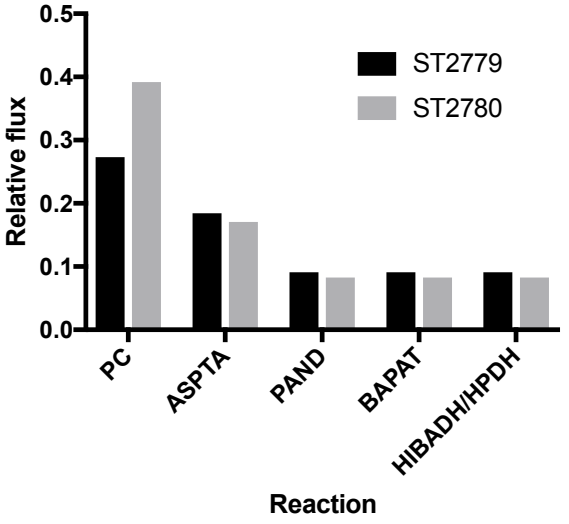


Figure S4. 3HP pathway fluxes from pyruvate, shown for both 3HP strains. Result of the ^{13}C-based metabolic flux analysis for the 3HP production pathway in the two strains tested. PC pyruvate carboxylase; ASPTA aspartate transaminase; PAND aspartate-1-decarboxylase; BAPAT β-alanine-pyruvate aminotransferase; HIBADH/HPDH 3-hydroxypropanoate dehydrogenase (NADH or NADPH dependent).

Table S1. List of primers used in this study

ID	Name	Sequence (5'->3')	Application
652	TcPAND_U1_fw	AGTGCAGGUAAAACAATGCCAGCTAC TGGTG	Forward primer for TcPanD biobrick
653	TcPAND_U1_rv	CGTGCGAUTCACAAATCGGAACCCAA TC	Reverse primer for TcPanD biobrick
5	pTEF1_fw<-	ACCTGCACUTTGTAATTAAAACTTAG	Forward primer for TEF promoter
6	pTEF1_rv<-	CACGCGAUGCACACACCATAGCTTC	Reverse primer for TEF promoter
8	pPGK1_rv->	ATGACAGAUTTGTTTTATATTTGTTG	Reverse primer for PGK1 promoter
7	pPGK1_fw->	CGTGCGAUGGAAGTACCTTCAAAGA	Forward primer for PGK1 promoter
11657	BcBAPAT_U1_fw<-	AGTGCAGGUAAAACAATGCACTTGCA AACTAAGACC	Forward primer for BcBAPAT
11658	BcBAPAT_U1_rv<-	CGTGCGAUTCACAATTGGGAGATGGA TTC	Reverse primer for BcBAPAT

704	EcYdfg_U2_fw->	ATCTGTCAUAAAACAATGATCGTTTTAGTAACTGGAG	Forward primer for EcYDFG
705	EcYdfg_U2_rv->	CACGCGAUTTACTGACGGTGGACATTC	Reverse primer for EcYDFG
670	ScPYC1_U1_fw<-	AGTGCAGGUAAAACAATGTCGCAAAGAAAATTCG	Forward primer for PYC1
671	ScPYC1_U1_rv<-	CGTGCGAUTCATGCCTTAGTTTCAACAG	Reverse primer for PYC1
672	ScPYC2_U2_fw->	ATCTGTCAUAAAACAATGAGCAGTAGCAAGAAATTG	Forward primer for PYC2
673	ScPYC2_U2_rv->	CACGCGAUTTACTTTTTTTGGGATGGG	Reverse primer for PYC2
82	pE2_NEW_fw->	ATCTGTCAUGGTACCAAAACAATG	Forward primer for PpHIBADH
12	pE2_rv->	CACGCGAUGTCGACTCA	Reverse primer for PpHIBADH
7011	ScMdh3deltaSKL_U1_fw<-	AGTGCAGGUAAAACAATGGTCAAAGTCGCAATTC	Forward primer for MDH3, without mitochondrial localisation sequence
7012	ScMdh3deltaSKL_U1_rv<-	CGTGCGAUTTAAGAGTCTAGGATGAAACTCTTG	Reverse primer for MDH3, without mitochondrial localisation sequence
278	pTDH3_fw->	CGTGCGAUATAAAAAACACGCTTTTTCAGTTCG	Forward primer for the promoter of TDH3
279	pTDH3_rv->	ATGACAGAUTTTGTTTGTTTATGTGTGTTTATTCGA	Reverse primer for the promoter of TDH3
1852	pTDH3_fw<-	CACGCGAUATAAAAAACACGCTTTTTCAG	Forward primer for the promoter of TDH3
1853	pTDH3_rv<-	ACCTGCACUTTTGTTTGTTTATGTGTGTTTATTC	Reverse primer for the promoter of TDH3
7029	SpMae1_U2_fw->	ATCTGTCAUAAAACAATGGGTGAACTCAAGGAAATC	Forward primer for SpMAE1
7030	SpMae1_U2_rv->	CACGCGAUTTAAACGCTTTCATGTTCAC	Reverse primer for SpMAE1
294	pTEF1_fw->	CGTGCGAUGCACACACCATAGCTTCAAAATG	Froward primer for the promoter of TEF1
295	pTEF1_rv->	ATGACAGAUTTGTAATTAAAACTTAGATTAGATTGCTATG	Reverse primer for the promoter of TEF1
905	X-4-up-out-sq	CTCACAAAGGGACGAATCCT	Primer for verification of integrations into site X-4
906	X-4-down-out-sq	GACGGTACGTTGACCAGAG	Primer for verification of integrations into site X-4
907	Ny-XI-1-up-sq	CTTAATGGGTAGTGCTTGACACG	Primer for verification of integrations into site XI-1
908	Ny-XI-1-down-sq	GAAGACCCATGGTTCCAAGGA	Primer for verification of integrations into site XI-1

901	X-2-up-out-sq	TGCGACAGAAGAAAGGGAAG	Primer for verification of integrations into site X-2
902	X-2-down-out-sq	GAGAACGAGAGGACCCAACAT	Primer for verification of integrations into site X-2
224	ADH1_test_fw	GAAATTCGCTTATTTAGAAGTGTC	A primer that together with CYC1_test_rv primer can be used to test cloning into yeast expression vectors
225	CYC1_test_rv	CTCCTTCCTTTTCGGTTAGAG	A primer that together with ADH1_test_fw primer can be used to test cloning into yeast expression vectors

Table S2. List of plasmids used in this study

Details of integration sites can be found in (Mikkelsen et al., 2012), except for the yeast retrotransposon site TY4, which can be found in (Maury et al., 2016).

ID	Name	Parental vector	Biobricks	Description	Source
pCfB799	pTY4-TcPanD<-PTEF1	pCfB322	tcPanD<-pTEF1	Integration into TY4 carrying TcPanD under influence of TEF1 promoter with Ura marker	(Borodina et al., 2015)
pCfB800	pX-4-loxP-SpHiS5-BcBAPAT<-PTEF1-PPGK1->EcYdfG	pCfB258	bcBAPAT<-pTEF1-pPGK1->ecYdfG	Integration into site X-4 of BcBAPAT under pTEF1 control, and EcYDFG under pPGK1 control with His marker	(Borodina et al., 2015)
pCfB743	pXI-1-loxP-KlLEU2-PYC1<-PTEF1-PPGK1->PYC2	pCfB358	PYC1<-pTEF1-pPGK1->PYC2	Integration into site XI-1 of PYC1 under pTEF1 control, and PYC2 under pPGK1 control with Leu marker	(Borodina et al., 2015)
pCfB801	pX-4-loxP-SpHiS5-BcBAPAT<-PTEF1-PPGK1->PpHIBADH	pCfB258	bcBAPAT<-pTEF1-pPGK1->ppHIBADH	Integration into site X-4 of bcBAPAT under pTEF1 control, and ppHIBADH under pPGK1 control with His marker	(Borodina et al., 2015)
pCfB2154	pX-2-loxP-ura3-ScMdh3deltaSKL-SpMae1	pCfB255	scMdh3deltaSKL<-pTDH3-pTEF1->spMae1	Integration into site X-2 of MDH3 lacking peroxisomal localisation sequence under	This study

				pTDH3 control, and Schizosaccharomyces pombe malate transporter MAE1 under pTEF1 control with Ura marker	
pCfB2226	X-4-loxP-SpHIS5syn	pCfB258	SpHIS5syn	Integration into site X-4 of His marker, used to close His auxotrophy	This study
pCfB7866	X-4pTDH3->lpLLDH	pCfB258	TDH3->lpLLDH	Integration into site X-4 of *Lactobacillus plantarum* L-LDH under pTDH3 control with URA marker	This study
pCfB322	pTY4-KlURA3-TAG	N/A	N/A	Backbone vector for multiple cassette integration into yeast retrotransposon site TY4	(Maury et al., 2016)
pCfB258	pX-4-LoxP-SpHiS5	N/A	N/A	Backbone vector for cassette integration into site X-4	(Jensen et al., 2014)
pCfB388	pXI-1-LoxP-LEU	N/A	N/A	Backbone vector for cassette integration into site XI-1	(Jensen et al., 2014)
pCfB255	pX-2-loxP-KlURA3	N/A	N/A	Backbone vector for cassette integration into site X-2	(Jensen et al., 2014)

Table S3. List of strains used in this study

ID	Name	Description	Plasmids	Source
ST001	CEN.PK113-7D	CEN.PK WT strain without auxotrophies	N/A	Petter Kötter
ST2606	TAM -Ura	Pdc- strain with uracil auxotroph used as the parental strain for ST7031	N/A	(van Maris et al., 2004a)
ST2681	TAM –Ura –His -Leu	Pdc- strain with uracil, histidine, and leucine auxotroph used as the parental strain for ST2779, 2780, and 5413	N/A	This study
ST2803	TAM	Pdc- strain with all markers closed, used for ^{13}C- MFA	N/A	This study
ST2779	3HP Producer – HPDH	3HP producing strain with the HPDH (NADPH dependent, from Escherichia coli) 3HP dehydrogenase	pCfB799 pCfB800 pCfB743	This study
ST2780	3HP Producer - HIBADH	3HP producing strain with the HIBADH from (NADH dependent, from Pseudomonas putida) 3HP dehydrogenase	pCfB799 pCfB801 pCfB743	This study
ST5413	Malate Producer	Malate producer, with native PYC1, PYC2, and MDH3 (lacking peroxisomal localisation	pCfB743 pCfB2154	This study

		tag) overexpressed, and *S. pombe* MAE1 expression.	pCfB2226	
ST7031	Lactate Producer	Lactate producer, with *L. plantarum* L-LDH expressed.	pCfB7866	This study

Table S4. Sequences of heterologous genes used in this study

For the sequences used in the 3HP producing strains, please refer to the supplementary materials found in (Borodina et al., 2015).

ID	Source
Lactobacillus plasntarum ldh1	AAAACAATGTCCTCTATGCCAAACCATCAAAAGGTTGTATTGGTTGGTGATGGTGCTGTTGGTTC TTCTTATGCTTTTGCTATGGCTCAACAAGGTATTGCCGAAGAATTTGTTATCGTTGATGTTGTCA AGGACAGAACTAAGGGTGATGCTTTGGATTTGGAAGATGCTCAAGCTTTTACTGCCCCAAAAAA GATCTATTCCGGTGAATACTCTGATTGCAAGGATGCTGATTTGGTTGTTATTACTGCTGGTGCTC CACAAAAACCAGGTGAATCAAGATTAGATTTGGTCAACAAGAACTTGAACATCTTGTCCTCAATC GTTAAGCCAGTTGTTGATTCTGGTTTCGATGGTATTTTCTTGGTTGCTGCTAACCCAGTTGATAT TTTGACTTACGCTACTTGGAAGTTCTCTGGTTTTCCAAAGGATAGAGTTATCGGTTCTGGTACTT CTTTGGACTCTTCTAGATTGAGAGTTGCTTTGGGTAAGCAATTCAATGTTGATCCAAGATCCGTT GATGCCTACATTATGGGTGAACATGGTGATTCTGAATTCGCTGCTTATTCTACTGCTACCATTGG TACTAGACCAGTTAGAGATGTTGCTAAAGAACAAGGTGTTTCCGATGAAGATTTGGCCAAGTTA GAAGATGGTGTTAGAAACAAGGCCTACGACATTATTAACTTGAAGGGTGCTACTTTCTACGGTA TTGGTACTGCTTTGATGAGAATTTCCAAGGCCATTTTGAGAGATGAAAACGCTGTTTTGCCAGTT GGTGCTTATATGGATGGTCAATACGGTTTGAACGATATCTACATTGGTACACCAGCTGTTATTGG TGGTACAGGTTTGAAGCAAATTATCGAATCTCCATTGTCCGCCGATGAATTGAAAAAGATGCAA GATTCTGCTGCCACCTTGAAAAAGGTTTTAAACGATGGTTTAGCCGAATTGGAAAACAAATGA
***Saccharomyces cerevisiae MDH3*ΔSKL**	AAAACAATGGTCAAAGTCGCAATTCTTGGCGCTTCTGGTGGCGTGGGACAACCGCTATCATTACT GCTAAAATTAAGCCCTTACGTTTCCGAGCTGGCGTTGTACGATATCCGAGCTGCGGAAGGCATT GGTAAGGATTTATCTCACATCAACACCAACTCAAGTTGTGTCGGTTATGATAAGGATAGTATTGA GAACACCTTGTCAAATGCTCAGGTGGTGCTAATACCGGCTGGTGTTCCCAGAAAGCCCGGTTTA ACTAGAGATGATTTGTTCAAGATGAACGCCGGTATTGTCAAAAGCCTGGTAACCGCTGTTGGAA AGTTCGCACCAAATGCGAGGATTTTAGTCATTTCAAACCCTGTAAACAGTTTGGTCCCTATTGCT GTGGAAACTTTGAAGAAAATGGGTAAGTTCAAACCTGGAAACGTTATGGGTGTGACGAACCTTG ACCTGGTACGTGCAGAAACCTTTTTGGTAGATTATTTGATGCTAAAAAACCCCAAAATTGGACAA GAACAAGACAAAACTACAATGCACAGAAAGGTCACTGTTATTGGGGGTCATTCAGGGGAAACCA TTATCCCAATAATCACCGACAAATCGCTGGTATTTCAACTTGATAAGCAGTACGAGCACTTCATTC ATAGGGTCCAGTTCGGAGGTGATGAAATTGTCAAAGCTAAACAGGGCGCCGGTTCCGCCACGTT GTCCATGGCGTTCGCGGGGGCCAAGTTTGCTGAAGAAGTTTTGAGGAGCTTCCATAATGAGAAA CCAGAAACGGAGTCACTTTCCGCATTCGTTTATTTACCAGGCTTAAAAAACGGTAAGAAAGCGCA

	GCAATTAGTTGGCGACAACTCTATTGAGTATTTTTCCTTGCCAATTGTTTTGAGAAATGGTAGCG TAGTATCCATCGATACCAGTGTTCTGGAAAAACTGTCTCCGAGAGAGGAACAACTCGTTAATACT GCGGTCAAAGAGCTACGCAAGAATATTGAAAAAGGCAAGAGTTTCATCCTAGACTCTTGA
Schizosaccharomyces pombe mae1	AAAACAATGGGTGAACTCAAGGAAATCTTGAAACAGAGGTATCATGAGTTGCTTGACTGGAATG TCAAAGCCCCTCATGTCCCTCTCAGTCAACGACTGAAGCATTTTACATGGTCTTGGTTTGCATGTA CTATGGCAACTGGTGGTGTTGGTTTGATTATTGGTTCTTTCCCCTTTCGATTTTATGGTCTTAATA CAATTGGCAAAATTGTTTATATTCTTCAAATCTTTTTGTTTTCTCTCTTTGGATCATGCATGCTTTT TCGCTTTATTAAATATCCTTCAACTATCAAGGATTCCTGGAACCATCATTTGGAAAAGCTTTTCAT TGCTACTTGTCTTCTTTCAATATCCACGTTCATCGACATGCTTGCCATATACGCCTATCCTGATACC GGCGAGTGGATGGTGTGGGTCATTCGAATCCTTTATTACATTTACGTTGCAGTATCCTTTATATA CTGCGTAATGGCTTTTTTTACAATTTTCAACAACCATGTATATACCATTGAAACCGCATCTCCTGC TTGGATTCTTCCTATTTTCCCTCCTATGATTTGTGGTGTCATTGCTGGCGCCGTCAATTCTACACA ACCCGCTCATCAATTAAAAAATATGGTTATCTTTGGTATCCTCTTTCAAGGACTTGGTTTTTGGGT TTATCTTTTACTGTTTGCCGTCAATGTCTTACGGTTTTTTACTGTAGGCCTGGCAAAACCCCAAGA TCGACCTGGTATGTTTATGTTTGTCGGTCCACCAGCTTTCTCAGGTTTGGCCTTAATTAATATTGC GCGTGGTGCTATGGGCAGTCGCCCTTATATTTTTGTTGGCGCCAACTCATCCGAGTATCTTGGTT TTGTTTCTACCTTTATGGCTATTTTTATTTGGGGTCTTGCTGCTTGGTGTTACTGTCTCGCCATGG TTAGCTTTTTAGCGGGCTTTTTCACTCGAGCCCCTCTCAAGTTTGCTTGTGGATGGTTTGCATTCA TTTTCCCCAACGTGGGTTTTGTTAATTGTACCATTGAGATAGGTAAAATGATAGATTCCAAAGCT TTCCAAATGTTTGGACATATCATTGGGGTCATTCTTTGTATTCAGTGGATCCTCCTAATGTATTTA ATGGTCCGTGCGTTTCTCGTCAATGATCTTTGCTATCCTGGCAAAGACGAAGATGCCCATCCTCC ACCAAAACCAAATACAGGTGTCCTTAACCCTACCTTCCCACCTGAAAAAGCACCTGCATCTTTGGA AAAAGTCGATACACATGTCACATCTACTGGTGGTGAATCGGATCCTCCTAGTAGTGAACATGAAA GCGTTTAA

Table S5. Flux calculations from INCA used to constrain a genome scale metabolic model of *S. cerevisiae*.

	ST2803		ST7031		ST5413		ST2779		ST2780	
Reaction Equation	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error
Glc.ext -> G6P	0.80945	0.18805	2.99605	0.71735	4.16395	1.36405	0.8062	0.1088	0.6736	0.0952
AcCOOH -> AcCOOH.ext	0.00905	0.00595	0.01895	0.01215	0.0103	0.0055	0.0069	0.0037	0.0056	0.003
Glyc -> Glyc.ext	0.5154	0.2584	2.11435	0.99645	4.07965	2.43555	0.45105	0.15965	0.37225	0.13135
0.459*Ala + 0.161*Arg + 0.102*Asn + 0.297*Asp + 0.007*Cys + 0.105*Gln + 0.302*Glu + 0.29*Gly + 0.066*His + 0.193*Ile + 0.296*Leu + 0.286*Lys + 0.051*Met + 0.134*Phe + 0.165*Pro + 0.185*Ser + 0.191*Thr + 0.028*Trp + 0.102*Tyr + 0.265*Val + 0.519*Glycogen + 0.023*Trehalose + 0.808*Mannan + 1.135*Glucan + 0.051*AMP + 0.05*CMP + 0.051*GMP + 0.067*UMP + 0.0030*DAMP + 0.0024*DCMP + 0.0024*DGMP + 0.0036*DTMP + 0.0066*TAG + 0.0007*ergosterol + 0.0015*zymosterol + 0.0006*PA + 0.0062*PC + 0.0046*PE + 0.0055*PI + 0.0017*PS + 0.4623*AcCoA.c -> Biomass	0.06835	0.01645	0.13005	0.01375	0.0961	0.0137	0.0602	0.0056	0.05255	0.00705
0*Pyr.c -> Pyr.mnt	0.51195	0.20945	0.1565	0.1565	0.0108	0	0.000114355	1.65E-07	0	0
0*Pyr.m -> Pyr.mnt	0.02585	0.00265	0.24445	0.24445	0.0299	0	0.00009959	4.1E-08	0.0026	0
Pyr.mnt -> Pyr.fix	0.5413	0.2065	0.46875	0.20025	0.0407	0	0.000213945	1.75E-07	0.0026	0
G6P <-> F6P	-0.09375	0.14405	-0.0836	0.1451	-0.14355	0.36825	-0.0345	0.053	-0.0239	0.0417

Table S5. Flux calculations from INCA used to constrain a genome scale metabolic model of S. cerevisiae.

	ST2803		ST7031		ST5413		ST2779		ST2780	
Reaction Equation	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error
F6P <-> FBP	0.38775	0.12145	1.7028	0.4702	2.5606	0.9089	0.4009	0.0731	0.3329	0.0605
FBP <-> DHAP + G3P	0.38775	0.12145	1.7028	0.4702	2.5606	0.9089	0.4009	0.0731	0.3329	0.0605
DHAP <-> G3P	-0.12865	0.15265	-0.40755	0.58175	-1.51235	1.57915	-0.0486	0.0947	-0.0383	0.0802
G3P -> PG3	0.55175	0.13295	2.4339	0.5401	2.33	0.7313	0.60595	0.07615	0.50505	0.06705
PG3 -> PEP	0.5384	0.1299	2.411	0.5394	2.28575	0.73145	0.59295	0.07505	0.4941	0.066
PEP -> Pyr.c	0.50245	0.12145	2.34065	0.53845	2.23375	0.73025	0.56125	0.07335	0.46655	0.06345
DHAP -> Glyc3P	0.51695	0.25845	2.11755	0.99655	4.08215	2.43555	0.45255	0.15965	0.37345	0.13125
Glyc3P -> Glyc	0.5154	0.2584	2.11435	0.99645	4.07965	2.43555	0.45105	0.15965	0.37225	0.13135
G6P -> CO2 + Ru5P	0.7492	0.2322	2.8035	0.7368	4.0405	1.4303	0.71665	0.11945	0.5949	0.1003
Ru5P <-> R5P	0.26905	0.07945	0.97315	0.24605	1.3762	0.4765	0.25685	0.04015	0.2136	0.0339
Ru5P <-> X5P	0.48055	0.15315	1.8296	0.4907	2.66395	0.95415	0.4597	0.0797	0.38115	0.06655
X5P <-> G3P + TKC2	0.48055	0.15315	1.8296	0.4907	2.66395	0.95415	0.4597	0.0797	0.38115	0.06655
F6P <-> E4P + TKC2	-0.232	0.0758	-0.89695	0.24495	-1.31895	0.47695	-0.222	0.0397	-0.1838	0.033
S7P <-> R5P + TKC2	-0.24845	0.07725	-0.93195	0.24565	-1.34485	0.47685	-0.2377	0.0398	-0.1973	0.0334
F6P <-> G3P + TAC3	-0.24845	0.07725	-0.93195	0.24565	-1.34485	0.47685	-0.2377	0.0398	-0.1973	0.0334
S7P <-> E4P + TAC3	0.24845	0.07725	0.93195	0.24565	1.34485	0.47685	0.2377	0.0398	0.1973	0.0334
Mal -> Pyr.m + CO2	0.00265	0.00265	0.0044	0.0044	0.09365	0.03545	0.0045	0.0045	0.00505	0.00505
Pyr.c + CO2 -> OAA.c	0.1928	0.0471	0.3666	0.0454	1.68765	0.64325	0.3092	0.0594	0.2545	0.0452
AcCHO -> AcCOOH	0.0055	0.0055	0.0118	0.0118	0.00505	0.00505	0.00355	0.00355	0.00285	0.00285
AcCHO -> AcCoA.c	0.05535	0.01625	0.1021	0.0248	0.02365	0.00785	0.0459	0.0081	0.04075	0.00805
Pyr.m -> AcCoA.m + CO2	0.19325	0.05045	0.51365	0.08985	0.47495	0.11495	0.1506	0.0225	0.1258	0.0215
OAA.m + AcCoA.m -> Cit	0.17	0.0435	0.46955	0.07875	0.38785	0.11155	0.1284	0.0189	0.10645	0.01755
Cit -> ICit	0.17	0.0435	0.46955	0.07875	0.38785	0.11155	0.1284	0.0189	0.10645	0.01755
ICit -> AKG + CO2	0.17	0.0435	0.46955	0.07875	0.38785	0.11155	0.1284	0.0189	0.10645	0.01755
AKG -> Suc + CO2	0.1009	0.0291	0.33785	0.07095	0.2896	0.1109	0.06725	0.01705	0.0531	0.0137
Suc <-> Fum	0.10135	0.02925	0.3387	0.071	0.2903	0.1109	0.0677	0.0171	0.05345	0.01365
Fum <-> Mal	0.1278	0.0345	0.3889	0.0738	0.3271	0.1106	0.09105	0.01765	0.07385	0.01485
Mal <-> OAA.m	0.12705	0.03435	0.38715	0.07335	0.2338	0.0995	0.0886	0.0174	0.0706	0.0147
PG3 + Glu -> Ser + AKG	0.01475	0.00525	0.0289	0.0099	0.04675	0.00775	0.014	0.0034	0.01205	0.00315
Ser <-> Gly + C1	-0.0008	0.0038	-0.0009	0.0094	0.02425	0.00535	0.00015	0.00295	0.00015	0.00255
Gly <-> CO2 + C1	0.02675	0.00755	0.04895	0.01165	0.0106	0.0044	0.02195	0.00385	0.0194	0.0038
Pyr.c -> Pyr.m	0.30485	0.07955	0.70425	0.11725	0.54815	0.11395	0.2415	0.0369	0.2048	0.0361
OAA.c <-> OAA.m	0.0435	0.0108	0.0832	0.0101	0.15275	0.03845	0.04105	0.00645	0.03685	0.00755
AcCoA.c <-> AcCoA.m	-0.00615	0.00805	-0.01175	0.01955	-0.0589	0.0109	-0.00643435	0.00626565	-0.00557355	0.00532645
AKG -> Glu	0.3663	0.0881	0.69775	0.07405	0.48565	0.06965	0.46115	0.06225	0.38565	0.05355
Glu -> Gln	0.062	0.0149	0.1179	0.0125	0.0871	0.0124	0.05455	0.00505	0.04765	0.00645
Glu -> Pro	0.0113	0.0027	0.02145	0.00225	0.01585	0.00225	0.00995	0.00095	0.00865	0.00115

Table S5. Flux calculations from INCA used to constrain a genome scale metabolic model of S. cerevisiae.

	ST2803		ST7031		ST5413		ST2779		ST2780	
Reaction Equation	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error
Glu + Glu + CO2 + Gln + Asp -> AKG + Arg + Glu + Fum	0.011	0.0026	0.0209	0.0022	0.0155	0.0022	0.0097	0.0009	0.00845	0.00115
OAA.c + Glu -> Asp + AKG	0.14945	0.03675	0.28315	0.03715	0.1538	0.0236	0.27005	0.05795	0.2185	0.0436
Asp + Gln -> Asn + Glu	0.00695	0.00165	0.0133	0.0014	0.0098	0.0014	0.00615	0.00055	0.00535	0.00075
Pyr.c + Glu -> Ala + AKG	0.01835	0.01835	0.03255	0.03255	0.01205	0.01205	0.01485	0.01485	0.0133	0.0133
Pyr.m + Glu -> Ala + AKG	0.01915	0.01915	0.03225	0.03225	0.03425	0.01505	0.0149	0.0149	0.01345	0.01345
Asp -> Thr	0.083	0.0214	0.15505	0.02695	0.0615	0.0123	0.06995	0.00935	0.0617	0.0099
Thr -> Gly + AcCHO	0.05645	0.01565	0.10355	0.02375	0.0252	0.0089	0.04645	0.00785	0.0411	0.0079
Ser + AcCoA.c + Asp + CO2 -> Cys + AcCOOH + Suc + CO2	0.00047 8435	0.00011 4935	0.00090 6935	0.00009 3065	0.00067 245	0.00009 581	0.00042 1335	0.00003 8995	0.00036 7735	0.00004 9565
Asp + C1 + AcCoA.c -> Met + AcCOOH	0.00345	0.00085	0.0066	0.0007	0.0049	0.0007	0.0031	0.0003	0.00265	0.00035
AKG + AcCoA.c + Glu + Glu -> Lys + CO2 + AKG + AKG	0.01955	0.00465	0.0372	0.0039	0.0275	0.0039	0.0172	0.0016	0.015	0.002
AcCoA.m + Pyr.m + Pyr.m + Glu -> Leu + CO2 + CO2 + AKG	0.02025	0.00485	0.0385	0.0041	0.02845	0.00405	0.01785	0.00165	0.01555	0.00205
Thr + Pyr.m + Glu -> Ile + CO2 + AKG	0.0132	0.0032	0.02505	0.00265	0.01855	0.00265	0.0116	0.0011	0.01015	0.00135
Pyr.m + Pyr.m + Glu -> Val + CO2 + AKG	0.01815	0.00435	0.03445	0.00365	0.02545	0.00365	0.01595	0.00145	0.0139	0.0019
PEP + PEP + E4P + Glu -> Phe + CO2 + AKG	0.0092	0.0022	0.01745	0.00185	0.01285	0.00185	0.00805	0.00075	0.00705	0.00095
PEP + PEP + E4P + Glu -> Tyr + CO2 + AKG	0.00695	0.00165	0.0133	0.0014	0.0098	0.0014	0.00615	0.00055	0.00535	0.00075
Ser + R5P + PEP + E4P + PEP + Gln -> Trp + CO2 + G3P + Pyr.c + Glu	0.00195	0.00045	0.00365	0.00035	0.0027	0.0004	0.00165	0.00015	0.0015	0.0002
R5P + C1 + Gln + Asp -> His + AKG + Fum	0.0045	0.0011	0.0086	0.0009	0.0063	0.0009	0.00395	0.00035	0.00345	0.00045
CO2 + Asp + R5P + Gln -> UMP + CO2 + Glu	0.0084	0.002	0.016	0.0017	0.0118	0.0017	0.0074	0.0007	0.00645	0.00085
UMP + Gln -> CMP + Glu	0.00355	0.00085	0.0068	0.0007	0.00505	0.00075	0.00315	0.00025	0.00275	0.00035
R5P + Gln + Gln + Gly + Asp + C1 + C1 + CO2 -> IMP + Glu + Glu + Fum	0.0074	0.0018	0.01405	0.00145	0.0104	0.0015	0.0065	0.0006	0.00565	0.00075
IMP + Asp -> AMP + Fum	0.0037	0.0009	0.00705	0.00075	0.00525	0.00075	0.0033	0.0003	0.0029	0.0004
IMP + Gln -> GMP + Glu	0.00365	0.00085	0.00695	0.00075	0.00515	0.00075	0.0032	0.0003	0.0028	0.0004
AMP -> DAMP	0.00024 605	0.00005 911	0.00046 8065	0.00004 9505	0.00034 5835	0.00004 9275	0.00021 6685	0.00002 0055	0.00018 912	0.00002 549
CMP -> DCMP	0.00016 4035	0.00003 9405	0.00031 204	0.00003 3	0.00023 0555	0.00003 2845	0.00014 446	0.00001 337	0.00012 608	0.00001 699
GMP -> DGMP	0.00016 4035	0.00003 9405	0.00031 204	0.00003 3	0.00023 0555	0.00003 2845	0.00014 446	0.00001 337	0.00012 608	0.00001 699
UMP -> DUMP	0.00024 605	0.00005 911	0.00046 8065	0.00004 9505	0.00034 5835	0.00004 9275	0.00021 6685	0.00002 0055	0.00018 912	0.00002 549
DUMP + C1 -> DTMP	0.00024 605	0.00005 911	0.00046 8065	0.00004 9505	0.00034 5835	0.00004 9275	0.00021 6685	0.00002 0055	0.00018 912	0.00002 549
Glyc3P -> PA	4.10085 E-05	9.8515E -06	0.00007 8011	0.00000 825	5.76385 E-05	8.2125E -06	3.61145 E-05	3.3425E -06	0.00003 152	0.00000 4249
Glyc3P + Ser -> PS	0.00011 619	2.79105 E-05	0.00022 1035	0.00002 3375	0.00016 331	0.00002 327	0.00010 2322	9.4685E -06	8.93045 E-05	1.20355 E-05
Glyc3P + Ser -> PE + CO2	0.00031 44	0.00007 553	0.00059 8085	0.00006 3255	0.00044 19	0.00006 296	0.00027 6875	0.00002 5625	0.00024 1655	0.00003 2575
Glyc3P + Ser + C1 + C1 + C1 -> PC + CO2	0.00042 376	0.00010 18	0.00080 611	8.525E-05	0.00059 56	8.486E-05	0.00037 318	0.00003 454	0.00032 571	0.00004 39
Glyc3P + G6P -> PI	0.00037 5915	0.00009 0305	0.00071 51	0.00007 563	0.00052 8355	0.00007 5275	0.00033 105	0.00003 064	0.00028 8935	0.00003 8945
Glyc3P -> TAG	0.00045 1095	0.00010 8365	0.00085 812	0.00009 075	0.00063 4025	0.00009 0335	0.00039 726	0.00003 677	0.00034 6725	0.00004 6735
AcCoA.c + AcCoA.c -> AcAcCoA	0.00089 273	0.00020 727	0.0017	0.0002	0.00125	0.00015	0.00079 4515	0.00007 3535	0.00069 344	0.00009 347
AcAcCoA + AcCoA.c -> HMGCoA	0.00089 273	0.00020 727	0.0017	0.0002	0.00125	0.00015	0.00079 4515	0.00007 3535	0.00069 344	0.00009 347

Table S5. Flux calculations from INCA used to constrain a genome scale metabolic model of S. cerevisiae.

	ST2803		ST7031		ST5413		ST2779		ST2780	
Reaction Equation	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error	Flux	Std Error
HMGCoA -> MVL	0.00089273	0.00020727	0.0017	0.0002	0.00125	0.00015	0.000794515	0.000073535	0.00069344	0.00009347
MVL -> IPPP + CO2	0.00089273	0.00020727	0.0017	0.0002	0.00125	0.00015	0.000794515	0.000073535	0.00069344	0.00009347
IPPP -> DMAPP	0.000300735	0.000072245	0.00057208	0.0000605	0.000422685	0.000060225	0.00026484	0.00002451	0.00023 1145	0.00003 1155
DMAPP + IPPP + IPPP + DMAPP + IPPP + IPPP -> CO2 + CO2 + CO2 + zymosterol	0.000150365	0.000036125	0.00028604	0.00003025	0.00021134	0.00003011	0.00013242	0.00001226	0.00011 5573	1.55775 E-05
zymosterol + C1 -> ergosterol	4.78435 E-05	1.14935 E-05	9.10135 E-05	9.6265E -06	0.000067245	0.000009581	4.21335 E-05	3.8995E -06	3.67735 E-05	4.9565E -06
G6P + G6P -> Trehalose	0.00155	0.00035	0.003	0.0003	0.0022	0.0003	0.0014	0.0001	0.0012	0.0002
G6P -> Glycogen	0.0355	0.0085	0.06745	0.00715	0.0499	0.0071	0.0312	0.0029	0.02725	0.00365
G6P -> Glucan	0.07755	0.01865	0.1476	0.0156	0.10905	0.01555	0.0683	0.0063	0.05965	0.00805
F6P -> Mannan	0.05525	0.01325	0.10505	0.01115	0.07765	0.01105	0.0486	0.0045	0.04245	0.00575
CO2.ext -> CO2	8.59935	2.31555	27.3867	7.378	44.06555	15.329255	9.26555	1.66225	7.5867	1.3627
CO2 -> CO2.out	9.6942	2.5672	31.3324	8.093	47.7698	16.657	10.2415	1.7874	8.39815	1.46195
GLC_113C -> Glc.ext	0.6856	0.1594	2.53175	0.60475	3.5224	1.1569	0.6848	0.0924	0.5723	0.0807
GLC_U13C -> Glc.ext	0.12355	0.02885	0.4646	0.1114	0.64035	0.20935	0.121	0.0165	0.1014	0.0145
Pyr.c -> Lac.c	-	-	1.23815	0.53785	-	-	-	-	-	-
Lac.c -> Lac.ext	-	-	1.23815	0.53785	-	-	-	-	-	-
OAA.c -> Mal.c	-	-	-	-	1.37955	0.63215	-	-	-	-
Mal.c -> Mal.ext	-	-	-	-	1.37955	0.63215	-	-	-	-
Asp -> Bala + CO2	-	-	-	-	-	-	0.1405	0.0571	0.10505	0.04165
Bala -> MalSA	-	-	-	-	-	-	0.1405	0.0571	0.10505	0.04165
MalSA -> 3HP	-	-	-	-	-	-	0.1405	0.0571	0.10505	0.04165
3HP -> 3HP.ext	-	-	-	-	-	-	0.1405	0.0571	0.10505	0.04165

References

Abbott, D.A., Zelle, R.M., Pronk, J.T., and van Maris, A.J.A. (2009). Metabolic engineering of Saccharomyces cerevisiae for production of carboxylic acids: current status and challenges. FEMS Yeast Res. *9*, 1123–1136.

Anderlei, T., Zang, W., Papaspyrou, M., and Büchs, J. (2004). Online respiration activity measurement (OTR, CTR, RQ) in shake flasks. Biochem. Eng. J. *17*, 187–194.

Borodina, I., Kildegaard, K.R., Jensen, N.B., Blicher, T.H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M.J., Rosenstand, I., et al. (2015). Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. Metab. Eng. *27*, 57–64.

Branduardi, P., Sauer, M., De Gioia, L., Zampella, G., Valli, M., Mattanovich, D., and Porro, D. (2006). Lactate production yield from engineered yeasts is dependent from the host background, the lactate dehydrogenase source and the lactate export. Microb. Cell Fact. *5*, 4.

Camarasa, C., Bidard, F., Bony, M., Barre, P., and Dequin, S. (2001). Characterization of Schizosaccharomyces pombe Malate Permease by Expression in Saccharomyces cerevisiae . Appl. Environ. Microbiol. *67*, 4144–4151.

Canelas, A.B., van Gulik, W.M., and Heijnen, J.J. (2008). Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions. Biotechnol. Bioeng. *100*, 734–743.

Cardoso, J.G.R., Jensen, K., Lieven, C., Lærke Hansen, A.S., Galkina, S., Beber, M., Özdemir, E., Herrgård, M.J., Redestig, H., and Sonnenschein, N. (2018). Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories. ACS Synth. Biol. *7*, 1163–1166.

Casal, M., Paiva, S., Queirós, O., and Soares-Silva, I. (2008). Transport of carboxylic acids in yeasts. FEMS Microbiol. Rev. *32*, 974–994.

Colombié, S., Dequin, S., and Sablayrolles, J.M. (2003). Control of lactate production by Saccharomyces cerevisiae expressing a bacterial LDH gene. Enzyme Microb. Technol. *33*, 38–46.

Dalwadi, M.P., King, J.R., and Minton, N.P. (2017). Multi-timescale analysis of a metabolic network in synthetic biology: a kinetic model for 3-hydroxypropionic acid production via beta-alanine. J. Math. Biol.

Flikweert, M.T., van der Zanden, L., Janssen, W.M.T.M., Yde Steensma, H., van Dijken, J.P., and Pronk, J.T. (1996). Pyruvate decarboxylase: An indispensable enzyme for growth of Saccharomyces cerevisiae on glucose. Yeast *12*, 247–257.

Flikweert, M.T., van Dijken, J.P., and Pronk, J.T. (1997). Metabolic responses of pyruvate decarboxylase-negative Saccharomyces cerevisiae to glucose excess. Appl. Environ. Microbiol. *63*, 3399–3404.

Flikweert, M.T., de Swaaf, M., van Dijken, J.P., and Pronk, J.T. (1999). Growth requirements of pyruvate-decarboxylase-negative Saccharomyces cerevisiae. FEMS Microbiol. Lett. *174*, 73–79.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. *2*, 31–34.

Gombert, A.K., Moreira dos Santos, M., Christensen, B., and Nielsen, J. (2001). Network Identification and Flux Quantification in the Central Metabolism of Saccharomyces cerevisiae under Different Conditions of Glucose Repression. J. Bacteriol. *183*, 1441–1451.

Hanko, E.K.R., Minton, N.P., and Malys, N. (2017). Characterisation of a 3-hydroxypropionic acid-inducible system from Pseudomonas putida for orthogonal gene expression control in Escherichia coli and Cupriavidus necator. Sci. Rep. *7*, 1724.

Hohmann, S. (1991). Characterization of PDC6, a third structural gene for pyruvate decarboxylase in Saccharomyces cerevisiae. J. Bacteriol. *173*, 7963–7969.

Ishida, N., Saitoh, S., Tokuhiro, K., Nagamori, E., Matsuyama, T., Kitamoto, K., and Takahashi, H. (2005). Efficient Production of l-Lactic Acid by Metabolically Engineered Saccharomyces cerevisiae with a Genome-Integrated l-Lactate Dehydrogenase Gene. Appl. Environ. Microbiol. *71*, 1964–1970.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., and Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. *14*, 238–248.

Kildegaard, K.R., Jensen, N.B., Schneider, K., Czarnotta, E., Ozdemir, E., Klein, T., Maury, J., Ebert, B.E., Christensen, H.B., Chen, Y., et al. (2016). Engineering and systems-level analysis of Saccharomyces cerevisiae for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway. Microb. Cell Fact. *15*, 53.

Kim, S., and Hahn, J.-S. (2015). Efficient production of 2,3-butanediol in Saccharomyces cerevisiae by eliminating ethanol and glycerol production and redox rebalancing. Metab. Eng. *31*, 94–101.

Kim, K.-S., Pelton, J.G., Inwood, W.B., Andersen, U., Kustu, S., and Wemmer, D.E. (2010). The Rut pathway for pyrimidine degradation: novel chemistry and toxicity problems. J. Bacteriol. *192*, 4089–4102.

Kumar, V., Ashok, S., and Park, S. (2013). Recent advances in biological production of 3-hydroxypropionic acid. Biotechnol. Adv. *31*, 945–961.

Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., et al. (2010). Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. *6*, 390.

Lui, C., and Lievense, J. (2005). Lactic acid producing yeast - US Patent application 20050112737.

Maaheimo, H., Fiaux, J., Cakar, Z.P., Bailey, J.E., Sauer, U., and Szyperski, T. (2001). Central carbon metabolism of Saccharomyces cerevisiae explored by biosynthetic fractional (13)C labeling of common amino acids. Eur. J. Biochem. *268*, 2464–2479.

Maris, A.J.A. va., Konings, W.N., Dijken, J.P. va., and Pronk, J.T. (2004). Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. Metab. Eng. *6*, 245–255.

van Maris, A.J.A., Geertman, J.-M.A., Vermeulen, A., Groothuizen, M.K., Winkler, A.A., Piper, M.D.W., van Dijken, J.P., and Pronk, J.T. (2004a). Directed Evolution of Pyruvate Decarboxylase-Negative Saccharomyces cerevisiae, Yielding a C2-Independent, Glucose-Tolerant, and Pyruvate-Hyperproducing Yeast. Appl. Environ. Microbiol. *70*, 159–166.

van Maris, A.J.A., Winkler, A.A., Porro, D., van Dijken, J.P., and Pronk, J.T. (2004b). Homofermentative Lactate Production Cannot Sustain Anaerobic Growth of Engineered Saccharomyces cerevisiae: Possible Consequence of Energy-Dependent Lactate Export. Appl. Environ. Microbiol. *70*, 2898–2905.

Maury, J., Germann, S.M., Baallal Jacobsen, S.A., Jensen, N.B., Kildegaard, K.R., Herrgård, M.J., Schneider, K., Koza, A., Forster, J., Nielsen, J., et al. (2016). EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in Saccharomyces cerevisiae. PLoS One *11*, e0150394.

McAlister-Henn, L., Steffan, J.S., Minard, K.I., and Anderson, S.L. (1995). Expression and Function of a

Mislocalized Form of Peroxisomal Malate Dehydrogenase (MDH3) in Yeast. J. Biol. Chem. *270*, 21220–21225.

Mikkelsen, M.D., Buron, L.D., Salomonsen, B., Olsen, C.E., Hansen, B.G., Mortensen, U.H., and Halkier, B.A. (2012). Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab. Eng. *14*, 104–111.

Mo, M.L., Palsson, B.O., and Herrgard, M.J. (2009). Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst. Biol. *3*, 37.

Nagamori, E., Shimizu, K., Fujita, H., Tokuhiro, K., Ishida, N., and Takahashi, H. (2013). Metabolic flux analysis of genetically engineered Saccharomyces cerevisiae that produces lactate under micro-aerobic conditions. Bioprocess Biosyst. Eng. *36*, 1261–1265.

Nevoigt, E. (2008). Progress in Metabolic Engineering of Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. *72*, 379–412.

Nour-Eldin, H.H., Geu-Flores, F., and Halkier, B.A. (2010). USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. Methods Mol. Biol. *643*, 185–200.

Oud, B., Flores, C.-L., Gancedo, C., Zhang, X., Trueheart, J., Daran, J.-M., Pronk, J.T., and van Maris, A.J.A. (2012). An internal deletion in MTH1 enables growth on glucose of pyruvate-decarboxylase negative, non-fermentative Saccharomyces cerevisiae. Microb. Cell Fact. *11*, 131.

Poskar, C.H., Huege, J., Krach, C., Franke, M., Shachar-Hill, Y., and Junker, B.H. (2012). iMS2Flux--a high-throughput processing tool for stable isotope labeled mass spectrometric data used for metabolic flux analysis. BMC Bioinformatics *13*, 295.

Pronk, J.T., Yde Steensma, H., and Van Dijken, J.P. (1996). Pyruvate metabolism in Saccharomyces cerevisiae. Yeast *12*, 1607–1633.

Rogers, J.K., and Church, G.M. (2016). Genetically encoded sensors enable real-time observation of metabolite production. Proc. Natl. Acad. Sci. *113*, 2388–2393.

Schmitt, H.D., and Zimmermann, F.K. (1982). Genetic analysis of the pyruvate decarboxylase reaction in yeast glycolysis. J. Bacteriol. *151*, 1146–1152.

Seeboth, P.G., Bohnsack, K., and Hollenberg, C.P. (1990). pdc1 Mutants of Saccharomyces cerevisiae Give

Evidence for an Additional Structural PDC Gene: Cloning of PDC5, a Gene Homologous to PDC1. J. Bacteriol. *172*, 678–685.

de Smidt, O., du Preez, J.C., and Albertyn, J. (2008). The alcohol dehydrogenases of Saccharomyces cerevisiae: a comprehensive review. FEMS Yeast Res. *8*, 967–978.

Wasylenko, T.M., and Stephanopoulos, G. (2015). Metabolomic and (13)C-Metabolic Flux Analysis of a Xylose-Consuming Saccharomyces cerevisiae Strain Expressing Xylose Isomerase. Biotechnol. Bioeng. *112*, 470–483.

Young, J.D. (2014). INCA: a computational platform for isotopically non-stationary metabolic flux analysis. Bioinformatics *30*, 1333–1335.

Zelle, R.M., de Hulster, E., van Winden, W.A., de Waard, P., Dijkema, C., Winkler, A.A., Geertman, J.-M.A., van Dijken, J.P., Pronk, J.T., and van Maris, A.J.A. (2008). Malic Acid Production by Saccharomyces cerevisiae: Engineering of Pyruvate Carboxylation, Oxaloacetate Reduction, and Malate Export . Appl. Environ. Microbiol. *74*, 2766–2777.

Research Article 3

Omic data based improvement of a high performance 3-hydroxypropionic acid producing strain of *Saccharomyces cerevisiae*

Mathew M Jessop-Fabre[1], Kanchana R Kildegaard[1], Birgitta E Ebert[2], Emre Özdemir[1], Lars M Blank[2], and Irina Borodina[1]. *Manuscript in preparation.* 2018.

[1] The Novo Nordisk Foundation for Biosustainability, Technical University of Denmark, Building 220, 2800 Kongens Lyngby, Denmark
[2] Institute of Applied Microbiology, RWTH Aachen University, Worringer Weg 1, 52056 Aachen, Germany

Abstract

Background

Bio-based production of fuels and chemicals must be realised if society is to be able to reduce its dependence on petroleum. 3-Hydroxypropionic acid (3HP) is one such chemical that has the opportunity to be produced *via* biological means. There is a high demand for 3HP derived products (the acrylic acid market alone is predicted to be worth USD 18.8 billion by 2020), but current biological methods of production are unable to reach the titres, rates, and yields required to make the process economically viable. Many tools have been developed for the rational prediction of metabolic engineering targets, but success of these tools is still limited.

Results

In this study, we apply -omics level approaches to provide a characterisation of a high 3HP producing strain of the yeast *S. cerevisiae* (ST938), shown previously to be capable of titres of 14 g L^{-1}. We first applied a transcriptomic analysis to ST938 to compare it to the WT strain. Introduction of the 3HP pathway caused widespread changes in the expression of genes across the genome. The engineered strain saw upregulation across glycolysis and the pentose phosphate pathway. We further applied ^{13}C-based metabolic flux analysis to investigate how flux distribution was altered through implementation of a high production 3HP pathway. The introduced pathway was shown to successfully divert flux from pyruvate, limiting the amount of flux into the TCA cycle. Significant flux was delivered to oxaloacetate, a precursor for 3HP biosynthesis, but only ~50% of this flux continues into 3HP suggesting that there is still space to improve the efficiency of the downstream pathway. Using a combination of metabolic flux analysis, genome scale metabolic modelling, and transcriptomic profiling, we predicted a number of gene deletion strategies for the improvement of 3HP production. We characterised three of these strategies *in vivo*

through shake flask fermentations. The deletion of a gene connected to acetylated sterol transport - *PRY1*, increased 3HP titres by ~27% compared to the parental strain, opening questions into the role of this gene on the production of 3HP, and whether similar effects may be seen in other production pathways. The two other deletion strategies (*JEN1* and *PGI1* deletion) did not improve 3HP production, but have highlighted new targets for further research.

Conclusions

In this study, we present data on the transcriptome and fluxome of a high-level 3HP producer. We use the generated data to predict metabolic engineering strategies. One of the predicted strategies resulted in a ~27% improvement in 3HP titres, while other strategies suggested future work directions. This demonstrates the value of –omic data analysis for strain engineering.

1 Introduction

Overuse of fossil fuels has led to a number of critical environmental issues that have left the global community with pressing challenges that must be addressed. Currently, many chemicals are derived from petroleum, and as such their long term use is generally not desirable. Wide arrays of chemicals can be sustainably produced from unicellular organisms such as yeasts and bacteria, and can help to reduce the global reliance on petrochemicals. 3-Hydroxypropionic acid (3HP) is one such chemical that can be produced biologically and can be used to generate derivative chemicals that have economic value. 3HP is a precursor of acrylic acid, and can also be polymerised to create a range of biodegradable and super-absorbent polymers. The market for acrylic acid is forecast to continue growing at a fast rate (compound annual growth rate of 7.6%), and is expected to be worth USD 18.8 billion by 2020 (Acrylic acid market report, Allied Market Research). Acrylic acid and 3HP are bulk chemicals, which means that they are produced in high amounts and therefore cheap to buy, meaning that they must be produced efficiently in order for bio-based production methods to compete with current petrochemical routes (Gustavsson and Lee, 2016; Shin et al., 2013).

3HP is a natural metabolic intermediate in many archaea, as well as in at least one species of bacteria (Alber et al., 2006; Hugler et al., 2002; Zarzycki et al., 2009). *Chloroflexus aurantiacus* is a marine, phototrophic bacteria capable of fixing CO_2 with 3HP as a metabolic intermediate, in what is known as the 3HP cycle (Holo, 1989; Tabita, 2009). It has also found to be a metabolic end product in some *Lactobacillus* species (Sobolov and Smiley, 1960; Talarico and Dobrogosz, 1990).

High 3HP titres from glucose have been reported in engineered bacteria, although industrial production in bacterial hosts can be hampered by the requirement for fermentations to be carried out at near neutral pH (Abbott et al., 2009; Liu et al., 2018; Song et al., 2016). Efficient recovery of acids in bioprocessing is often dependent on the acid existing in its undissociated form, which predominates at pH values below the pKa

values of the acid (pKa for 3HP is 4.2). Downstream processing costs can therefore be minimised by performing fermentations at low pH, making organisms that can natively tolerate low pH such as *Saccharomyces cerevisiae,* attractive hosts for industrial production of acids (Van Maris et al., 2004).

Two biosynthetic pathways have been shown to have promise for 3HP production from glucose in *S. cerevisiae*. One of these pathways utilises malonyl-CoA as a precursor to 3HP (Kumar et al., 2013). The malonyl-CoA reductase (MCR) gene from *C. aurantiacus* has been shown to be functional in *S. cerevisiae*, catalysing the two step reaction of malonyl-CoA into 3HP *via* malonic semialdehyde (Alber et al., 2006; Hugler et al., 2002; Kildegaard et al., 2016). Natural cytoplasmic levels of acetyl-CoA, the precursor to malonyl-CoA, in *S. cerevisiae* are low and tightly regulated under glucose abundant conditions, with much research focussing on increasing cytosolic pools for the production of acetyl-CoA derived products (Chen et al., 2013; Kozak et al., 2014; Liu et al., 2017; Lv et al., 2016; Nielsen, 2014; Shiba et al., 2007). Native acetyl-CoA biosynthesis is dependent on ATP and therefore high levels of aeration are required for maximal production (Pronk et al., 1996; van Rossum et al., 2016).

The second pathway passes through β-alanine as an intermediate metabolite. Borodina *et al.*, (2015) demonstrated that this pathway is theoretically capable of higher yields than the malonyl-CoA pathway, and it is not as dependent on high levels of oxygenation during fermentation, making the process more economically attractive. Borodina *et al.*, (2015) tested this pathway *in vivo* in *S. cerevisiae*, and showed that β-alanine can be produced to high levels in yeast through the introduction of a heterologous aspartate-1-decarboxylase from *Tribolium castaneum*, *panD*. Pyruvate and β-alanine are interconverted to L-alanine and malonic semialdehyde *via* expression of *yhxA* from *Bacillus cereus*. In the final enzymatic step, *Escherichia coli ydfG* reduces malonic semialdehyde to 3HP (with NADPH as cofactor). The final optimised strain (ST938) produced 13.7 ± 0.3 g L^{-1} 3HP in a controlled fed-batch fermentation, representing a 14% C-mol/C-mol yield on glucose.

While these results show that *S. cerevisiae* is a promising cell factory for 3HP production, further characterisation is required to enable engineering efforts to reach titres, rates, and yields that are required to make the process economically relevant.

The "-omic" technologies are routinely used for strain characterisation in metabolic engineering practices to identify target areas for strain improvement (Petzold et al., 2015). In the present study we present the use of transcriptomic and ^{13}C-based metabolic flux analysis (MFA) for the characterisation of ST938. We used these data to constrain a genome scale metabolic model of ST938 and to predict metabolic engineering strategies for achieving improvements in 3HP production.

2 Materials and methods

2.1 Strains

E. coli DH5α was used for cloning, plasmid propagation, and plasmid storage. The parental *S. cerevisiae* strain used as the chassis for strain construction was an auxotrophic variant of CEN.PK (MATa ura- his- leu-), obtained as a kind gift from Peter Kötter, Johann Wolfgang Goethe University, Frankfurt, Germany. The engineered 3HP producing strain, ST938, is taken from the previously published work of (Borodina et al., 2015).

2.2 Knock-out cassette construction

For gene knock-outs, kanMX expression cassettes were amplified from the yeast knock-out collection for each target knock-out (Giaever et al., 2002). Each strain from the collection containing the desired knock-outs was re-streaked from glycerol stocks onto YPD agar plates containing 200 mg L^{-1} G418. Individual colonies were picked and served as the template for the amplification of the desired knock-out cassettes. Primers for knock-out cassette amplification were chosen according to the accompanying materials of the knock-out collection. PCR was performed on each knock-out strain, and the successfully

amplified cassettes were gel purified before transformation into ST938. Complete lists of knock-out cassettes, and primers are included in the supplementary materials.

2.3 Yeast transformations

Yeast were chemically transformed with 500 ng of purified gene knock-out cassettes according to the lithium acetate method (Gietz and Schiestl, 2007). All transformants were selected on synthetic complete agar plates lacking uracil, histidine, and leucine to maintain selection pressure in ST938. 200 mg L^{-1} G418 was added to the synthetic complete (SC) agar plates for selection of the gene knock-out cassettes. The agar plates were incubated at 30 °C for 2-3 days, after which time individual colonies were tested by PCR to confirm the correct integration of DNA fragment. Primers used for confirmation of insertions can be found in the supplementary materials. Correct isolates were then saved as glycerol stocks at -80 °C.

2.4 Fermentation conditions and media

For propagation of the strains from the knock-out collection, either YPD agar, or YPD broth containing 200 mg L^{-1} G418 was used. YPD was prepared from 20 g L^{-1} glucose, 20 g L^{-1} bacteriological peptone, 10 g L^{-1} yeast extract, and for agar plates; 10 g L^{-1} agar was added. For propagation of ST938, SC media lacking uracil, leucine, and histidine was prepared from SC drop-out powders. Selection and propagation of the knock-out variants of ST938 was performed with SC agar plates or SC liquid media, lacking uracil, leucine, and histidine, as well as containing 200 mg L^{-1} G418. Mineral media (MM) for the fermentations of the strains was prepared as previously reported, but with 50 g L^{-1} final glucose concentration in the 96-deep well plate screening experiment, and 100 g L^{-1} in the shake flask fermentation experiment (Jensen et al., 2014).

Small-scale fermentations were performed in 96-deep well plates, sealed with gas permeable membranes (EnzyScreen, Germany). Each strain was inoculated from glycerol stock into a 12 mL pre-culture tube containing 5 mL SC media (lacking leucine, uracil, and histidine and with G418 supplementation) and left overnight in a shaking incubator set to 30 ˚C and 250 rpm. The following day, each well was filled with 500 μL of MM, and inoculated with the strain pre-cultures to a starting OD_{600} of 0.05. The 96-deep well plates were placed in a shaking incubator (250 rpm, 30 ˚C) for 72 hours, after which time samples were taken for analysis. These samples were measured for cellular density at 600 nm on an Implen P300 spectrophotometer, spun down at 10,000 x g for 10 minutes, and the supernatant stored at -20 ˚C until HPLC analysis of 3HP and selected metabolites was performed. Each fermentation was carried out in quadruplicate.

For the shake flask fermentations, each strain was pre-cultured as for the 96-deep well plate fermentations. The following evening the strain cultures were inoculated into baffled 250 mL shake flasks containing 25 mL of MM with 100 g L^{-1} glucose to a starting OD_{600} of 0.05. The shake flasks were placed in a shaking incubator (300 rpm, 30 ˚C) for a total of 96 hours with samples being taken at regular intervals and analysed as described above. Each fermentation was carried out in triplicate.

2.5 Analytical methods

Quantification of 3HP, glucose, and secreted metabolites was performed *via* HPLC. Samples for analysis were taken, centrifuged, and the supernatant transferred to Nunc 96-well plates with rubber sealed tops (Thermo Scientific, USA). Samples were separated in an Aminex HPX-87H ion exclusion column (Bio-Rad, USA), with a column temperature of 60 ˚C and flow rate of 600 μL min^{-1}. A Dionex RI-101 Refractive Index Detector, and a DAD-3000 Diode Array Detector at 45 ˚C measured the compounds after separation (Dionex, USA). 30 μL of each sample was run for 45 minutes, with 1 mM H_2SO_4 as the eluent.

2.6 Chemicals and reagents

Unless otherwise stated, all reagents and chemicals were purchased from Sigma-Aldrich, USA.

2.7 Transcriptomic and gene set enrichment analysis

The reference WT strain, and strain ST938 were grown in 1 L continuous cultivation bioreactors. After three residence times, biomass samples were harvested and total RNA extraction performed as previously described in (Kildegaard et al., 2014). Sequencing library preparation and cDNA sequencing was performed also as previously described in (Kildegaard et al., 2016). Differential gene expression analysis was performed with TopHat (2.0.13), and Cufflinks (2.2.1) with the reference genome for CEN.PK113-7D (Trapnell et al., 2012). Gene set enrichment analysis was performed with the goseq package for R, taking into account gene length bias (Young et al., 2010). The false discovery rate threshold for considering significantly enriched GO terms was set to < 0.05, following the method of Benjamini and Hochberg (1995).

2.8 ^{13}C-labelled fermentations and flux analysis

The labelled glucose fermentations, sample preparation, GC-MS analysis, and metabolic flux analysis were performed as previously published (Kildegaard et al., 2016). Cultivations were performed in 1 L bioreactors. Experimental fluxes were estimated with the FiatFlux toolbox for MATLAB (Zamboni et al., 2005). The calculated fluxes were used to constrain the iMM904 genome scale metabolic model of *S. cerevisiae* (Mo et al., 2009). Beneficial gene deletion strategies were predicted through the OptKnock software package implemented in the the CAMEO toolbox for Python (Burgard et al., 2003; Cardoso et al., 2018).

2.10 Statistics

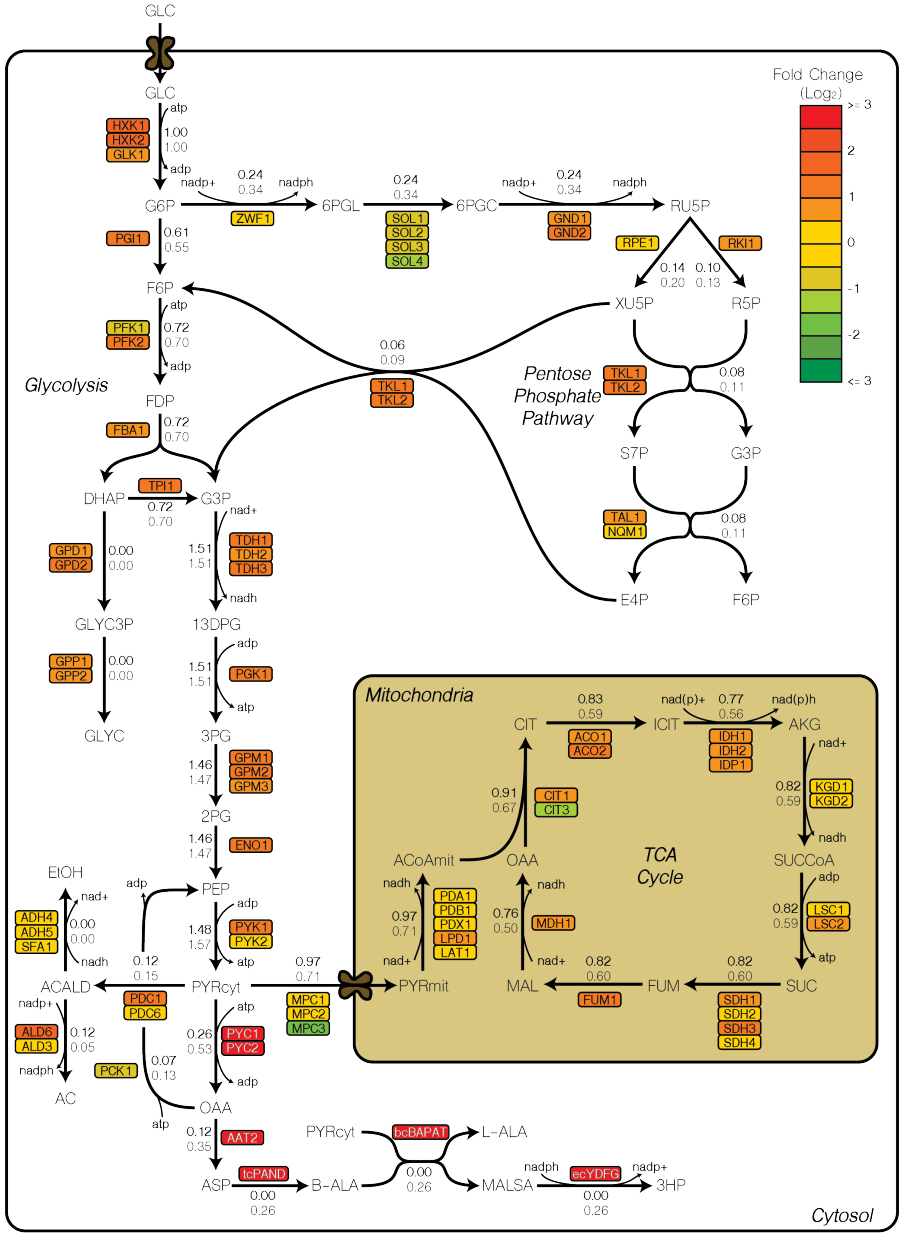
The data generated from the shake flask fermentations were analysed with the GraphPad Prism software version 7 (GraphPad Software, Inc.). Statistical significance was calculated with Dunnett's multiple comparisons test for one-way ANOVA. All values are expressed as the mean average ± SD.

3 Results and discussion

3.1 Transcriptome analysis

To analyse how high-level production of 3HP affects the physiology of *S. cerevisiae*, we analysed the transcriptomic profile of the high-producer strain (ST938) against the WT

Figure 1 (See over page). Map of metabolic fluxes and transcriptomic profile of the central carbon metabolism of ST938 compared to the WT strain (CEN.PK). Numbers over the reactions are the glucose uptake normalised fluxes, with the bold WT value on top, and the engineered strain value on the bottom. Key reactions were measured through ^{13}C metabolic flux analysis, and the remaining reactions were calculated by fitting a genome scale metabolic model with the experimentally determined fluxes. Next to each reaction are the relative transcript levels of the main genes involved in each reaction. Genes written in white are heterologous or overexpressed genes. Relative transcript levels are taken as the log_2 of the ST938/WT levels. Red values signify genes that are more highly expressed in ST938, and green values are those more highly expressed in the WT strain. Abbreviations: GLC glucose, G6P glucose-6-phosphate, 6PGL D-6-phospho-glucono-δ-lactone, 6PGC 6-phospho-D-gluconate, RU5P ribulose-5-phosphate, XU5P xylulose-5-phosphate, R5P ribose-5-phosphate, S7P sedoheptulose-7-phosphate, G3P glyceraldehyde-3-phosphate, E4P erythrose-4-phosphate, F6P fructose-6-phosphate, FDP fructose-1,6-diphosphate, DHAP dihydroxy-acetone-phosphate, GLYC3P glycerol-3-phosphate, GLYC glycerol, 13DPG 1,3-diphosphateglycerate, 3PG 3-phosphoglycerate, 2PG 2-phosphoglycerate, PEP phosphoenolpyruvate, PYRcyt pyruvate (cytosolic), ACALD acetaldehyde, EtOH ethanol, AC acetate, PYRmit pyruvate (mitochondrial), ACoAmit acetyl-CoA (mitochondrial), CIT citrate, ICIT isocitrate, AKG α-ketoglutarate, SUCCoA succinyl-CoA, SUC succinate, FUM fumarate, MAL malate, OAA oxaloacetate, ASP aspartate, B-ALA β-alanine, L-ALA L-alanine, MALSA malonic semialdehyde, 3HP 3-hydroxypropionate.


GLC
Fold Change (Log2)
>= 3
2
1
0
-1
-2
<= 3
Glycolysis
Pentose Phosphate Pathway
Mitochondria
TCA Cycle
Cytosol
G6P
6PGL
6PGC
RU5P
F6P
XU5P
R5P
FDP
S7P
G3P
E4P
DHAP
GLYC3P
13DPG
GLYC
3PG
2PG
PEP
EtOH
ACALD
AC
PYRcyt
OAA
ASP
B-ALA
L-ALA
MALSA
3HP
CIT
ICIT
AKG
SUCCoA
SUC
FUM
MAL
OAA
ACoAmit
PYRmit

parental strain (CEN.PK). These data were generated from samples taken from carbon-limited continuous cultivations of the WT and ST938 strains. In total 6506 transcripts were measured, of which there were 1246 transcripts that were found to be significantly differentially expressed (to a q value < 0.05) between ST938 and the parental WT strain.

As shown in Figure 1, much of the central carbon metabolism pathways appear to be upregulated in ST938, particularly the reactions involved in glycolysis. ST938 exhibits a higher glucose uptake rate than the WT strain (0.659 *vs.* 0.463 $mmol^{-1}$ cDW^{-1} h^{-1}), and so this may cause genes involved in glucose assimilation to be upregulated in response. The hexokinases have a high degree of flux control over glycolysis, and their overexpression may then greatly increase the metabolism of glucose (Smallbone et al., 2013) *UBC8*, a gene involved in the negative regulation of gluconeogenesis, is significantly down-regulated in ST938, suggesting that this strain is under lower levels of glucose repression than WT. Most of the genes related to gluconeogenesis are also upregulated, although those involved in the glyoxylate cycle were not.

Several degradation of allantoin (*DAL)* genes were significantly upregulated in ST938, with *DAL2-6* showing higher than WT levels. The related degradation of urea gene *DUR3*, is also highly upregulated in ST938, suggesting that this strain is utilising the available nitrogen in the media faster than the WT, with the lower free nitrogen levels relieving nitrogen catabolite repression. Genes related to proline utilisation (*PUT1*, *PUT2*) were also found to be significantly upregulated compared to WT further indicating that the additional protein burden on ST938 may contribute to nitrogen starvation responses.

There were only two annotated membrane transporters that were upregulated in ST938, one of them being *MCH5. MCH5* encodes a plasma membrane transporter known to facilitate riboflavin transport in *S. cerevisiae* (Reihl and Stolz, 2005). The monocarboxylate

transporter-homologous (Mch) family of proteins have been reported to not transport monocarboxylic acids in yeast, although these results are those based on the uptake of lactate, pyruvate, and acetate, for which there exist other transporters such as Jen1p (Casal et al., 2008; Makuc et al., 2001). Further research would be needed to confirm that the Mch family are not involved in 3HP transport, and is only functioning to increase riboflavin uptake. Jen1p is the main transporter for pyruvate and lactate, and represents a candidate for 3HP transport due to 3HP's chemical similarity to both these, although transcript levels in this study were not found to be significantly different between the WT and ST938 strains. Genes related to acid stress such as; *YAP1, MSN2, MSN4,* and *PDR12,* did not show significant differences between the tested strains. Due to the cultivations being performed in a pH controlled bioreactor, it is expected that the strains should not experience significant acid stress.

We performed a gene set enrichment analysis on ST938 using the gene ontology (GO) goseq package for R (Young et al., 2010). Gene set enrichment analysis can highlight functions that are over- or under-represented in a transcriptional data set, generating a high-level understanding of the regulatory processes within the host cell. Our analysis of ST938 is shown in Figure 2, with the GO identifiers referring to biological processes, cellular component, and molecular function terms. Many GO terms are significantly downregulated in ST938 *vs.* the WT strain, with many of these terms being related to cellular maintenance, such as ribosome processes, peptide biosynthesis and translation, and RNA processing. This apparent downregulation of protein production is perhaps linked to the introduction of the heterologous pathway. In particular, *tcPanD* is introduced into the yeast Ty4 retrotransposon site and is therefore present in multiple copies. High protein expression may cause metabolic stress to the host cell, thereby reducing the efficiency of production (Dürrschmid et al., 2008; Haddadin and Harcum, 2005; Hoffmann and Rinas, 2004).

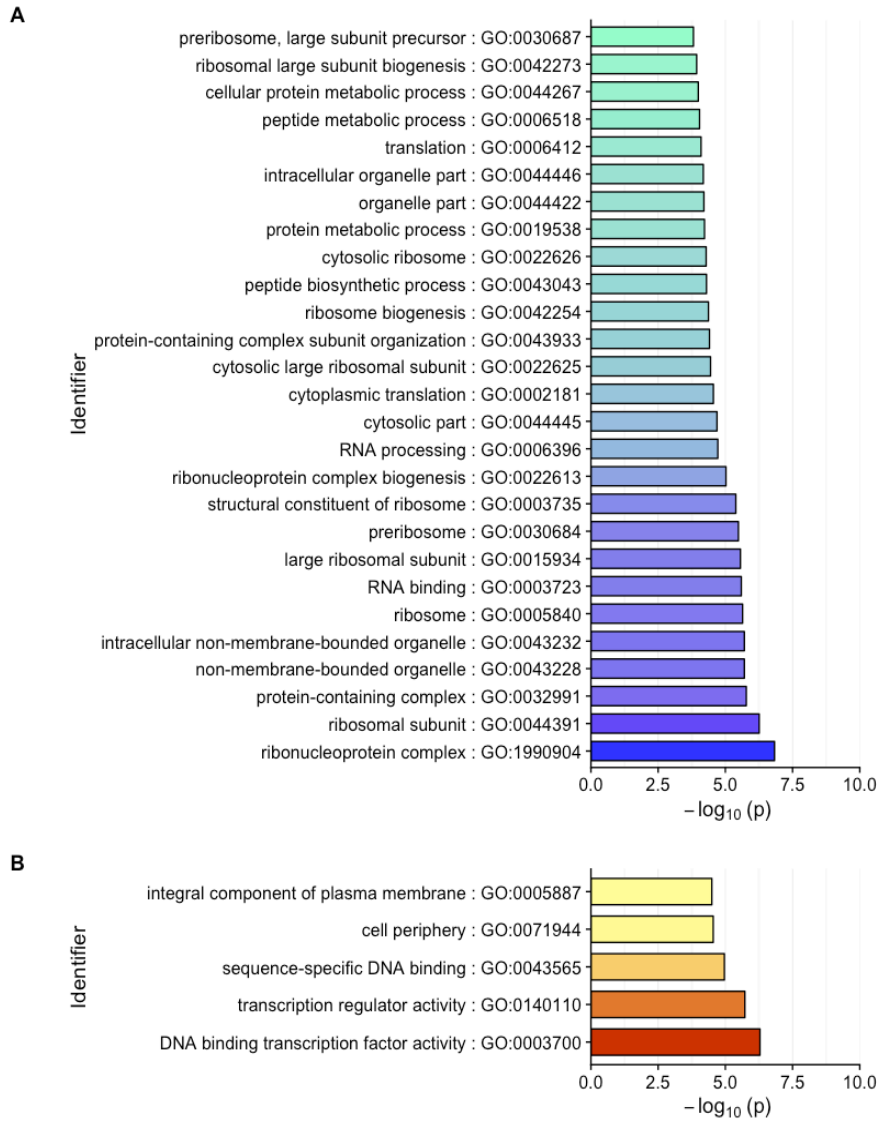


Figure 2. Gene ontology enrichment analysis of ST938. The transcriptomic profiles of the WT strain and strain ST938 were used to generate gene sets that were under-represented (A), and over-represented (B) compared to the WT. Shown are all categories that had a false discovery rate < 0.05. The x axis shows the $-\log_{10}$ of the p values, where a higher $-\log_{10}(p)$ value indicates a category with a higher statistical probability of being significant.

3.2 ^{13}C-based metabolic flux analysis

While informative, transcriptomic data is unable to provide information on metabolic fluxes. Therefore, we examined the metabolism of ST938 through ^{13}C-MFA. The experimentally derived fluxes were used to constrain the iMM907 genome-scale metabolic model (GEM) of *S. cerevisiae* (Mo et al., 2009), which we used to predict fluxes throughout the reaction network of ST938 (Figure 1). Our analysis shows that there are large changes between ST938 and the WT strain in the PP pathway, the TCA cycle, and the first steps of gluconeogenesis. The flux ratios through upper and lower glycolysis are close to identical in the WT strain and ST938. However, perhaps due to the increased uptake of glucose in ST938 (0.463 *vs.* 0.659 $mmol^{-1}$ cDW^{-1} h^{-1}), more flux gets diverted into the PP pathway to cope with the additional carbon flux.

Higher flux is carried through pyruvate decarboxylase in ST938, although it produces less acetate than the WT strain. The additional flux is predicted to be channelled into acetoin biosynthesis. This observation is backed by the transcriptomic data, where the gene involved in the conversion of acetoin to butanediol, *BDH1*, is significantly up-regulated in ST938.

The TCA cycle carries relatively less flux in ST938, compared to the WT strain. This effect is to be expected with the overexpression of pyruvate carboxylase genes. The resultant pool of oxaloacetate, cannot be fully converted to 3HP, and some of the excess flux is diverted through the phosphoenolpyruvate carboxykinase - Pck1p. The transcript levels of *PCK1*, are on the other hand, significantly lower in ST938 than in the WT. *PCK1* is known to be regulated by Mcm1p and Cat8p, and while the levels of *MCM1* mRNA remained almost the same between ST938 and the WT, *CAT8* mRNA levels were significantly lower in ST938. *CAT8* is normally expressed during the shift from

fermentative to oxidative metabolism (diauxic shift), and de-represses a range of genes involved in gluconeogenesis and ethanol utilisation (Haurie et al., 2001).

3.3 Data-guided modification of ST938

After analysis of both the transcriptomic and fluxomic data, we identified genes of interest that could be related to the performance of the 3HP pathway. We used the OptKnock algorithm on the ^{13}C-MFA constrained GEM of yeast to predict gene deletion strategies, along with insights from the transcriptomic data set (Burgard et al., 2003). The OptKnock algorithm predicted four different gene deletion strategies; *PSD1* – encoding a mitochondrially located phosphatidylserine decarboxylase, *PGI1* – encoding a phosphoglucose isomerase responsible for an early step in glycolysis, *OSM1* – encoding two forms of fumarate reductase, one locating to the mitochondria and the other to the endoplasmic reticulum, and *JEN1* – encoding a monocarboxylate transporter (Aguilera and Zimmermann, 1986; Casal et al., 1999; Clancey et al., 1993; Muratsubaki and Enomoto, 1998). An additional knockout strategy was tested based on a highly overexpressed transcript in ST938, *PRY1* – a gene encoding a protein implicated in acetylated sterol export (Choudhary and Schneiter, 2012).

3.4 Fermentation profiles of knock-out designs

After knock-out strains had been successfully constructed from the ST938 host *via* the disruption of the target genes by insertion of a kanMX cassette, they were screened for 3HP production *via* microtitre plate fermentations. Out of the five gene deletion strategies that were tested, two of these (*PSD1* and *OSM1*) showed only minimal 3HP production under microtitre plate fermentations (see Figure S3). The remaining three knock-out strains were further tested in shake flask fermentations in batch media containing 100 g L-

[1] glucose. Each fermentation profile can be found in Figure 3, and a list of fermentation characteristics can be found in Table 1.

The *JEN1* knockout strain, ST8717, is predicted to increase 3HP production by the OptKnock algorithm as this gene is linked to reactions that transport lactate and pyruvate out of the cell, and thus the deletion of it should redirect flux into 3HP production. Jen1p is a monocarboxylate transporter, and its deletion completely prevents the uptake of both lactate and pyruvate in a WT strain suggesting that this is the only functional transporter for the uptake of both these monocarboxylates (Akita et al., 2000; Casal et al., 1999). Due to the close chemical similarity between lactate and 3HP, Jen1p is a candidate for 3HP transport in yeast. If Jen1p is involved in 3HP export then it would be expected that a knock-out would reduce 3HP titres, and an intracellular build up of 3HP would cause the host to be growth impaired. ST8717, showed only a slightly reduced growth rate compared

*Table 1. Fermentation characteristics of strains used in this study. Three gene deletion strategies were successfully implemented in the parental ST938 strain. Each value is the mean average with standard deviation of shake flask experiments performed in triplicate. Each value is evaluated for statistical significance compared to the control group (ST938). **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, ns = non significant.*

	ST938	**Δ*PGI1* (ST8714)**	**Δ*JEN1* (ST8717)**	**Δ*PRY1* (ST8718)**
Growth Rate (h^{-1})	0.243 ± 0.004	0.198 ± 0.004 ****	0.231 ± 0.001 **	0.234 ± 0.004 *
Final OD_{600}	66.33 ± 1.53	41.00 ± 1.73 ****	68.33 ± 2.31 ns	61.33 ± 1.53 *
3HP (g L^{-1})	1.21 ± 0.06	0.68 ± 0.05 ****	1.15 ± 0.04 ns	1.54 ± 0.06 ***
Peak ethanol (g L^{-1})	32.05 ± 1.14	35.06 ± 2.03 ns	37.79 ± 1.26 **	34.14 ± 1.75 ns
Glycerol (g L^{-1})	3.67 ± 0.19	7.03 ± 0.61 ****	4.53 ± 0.20 ns	2.92 ± 0.34 ns
Acetate (g L^{-1})	5.95 ± 0.36	4.21 ± 0.26 **	5.53 ± 0.30 ns	6.33 ± 0.53 ns

to ST938 (0.231 h^{-1} ± 0.001 *vs.* 0.246 h^{-1} ± 0.004). At the end of 96 hour shake flask fermentations containing 100 g L^{-1} glucose, the final optical density is also very close to the parental strain ST938, suggesting that this modification is only slightly detrimental to cellular growth and biomass accumulation. However, final 3HP titres were not significantly different in ST8717 (1.15 ± 0.04 *vs.* 1.21 ± 0.06 g L^{-1}) indicating that this is not a successful strategy to increase 3HP production. This result does, however, suggest that Jen1p is not responsible for the majority of 3HP export in *S. cerevisiae* as there was no significant drop in 3HP titre which is not what might be expected if Jen1p were the major transporter for this molecule. Glycerol and acetate production did not appear to be significantly influenced by this gene deletion, but maximum ethanol titres were slightly higher than ST938. Further work is needed to identify the 3HP export mechanisms in *S. cereviasiae*, which may then lead to targets for improving export rates, in turn improving the production capabilities of the host strain. Several carboxylate transporters have been identified in *S. cerevisiae* and other yeasts, and further research may elucidate one that are optimal for 3HP transport (Casal et al., 2008).

A deletion of the phosphoglucose isomerase gene *PGI1* appears as an interesting strategy for increasing 3HP production in ST938. Pgi1p catalyses the interconversion of glucose-6-phosphate into fructose-6-phosphate, feeding into upper glycolysis. Disruption of this reaction diverts flux through the PP pathway (Fiaux et al., 2003). However, a deletion of *PGI1* renders *S. cerevisiae* incapable of growth on glucose as the sole carbon source (Aguilera, 1986). The cause for this is growth defect is thought to be the exhaustion of $NADP^+$ through the PP pathway that reduces 2 moles of $NADP^+$ per mole of glucose (Boles et al., 1993). The expression of a heterologous transhydrogenase from *E. coli*, UdhA, can restore the ability of a *PGI1* deletion mutant to grow on on glucose – albeit at a decreased growth rate of 0.15 h^{-1} (Fiaux et al., 2003). The final step in the 3HP production pathway of ST938 is the conversion of malonic semialdehyde to 3HP with the coupled generation of 2 moles of $NADP^+$ per mole of glucose. 3HP production *via* this pathway

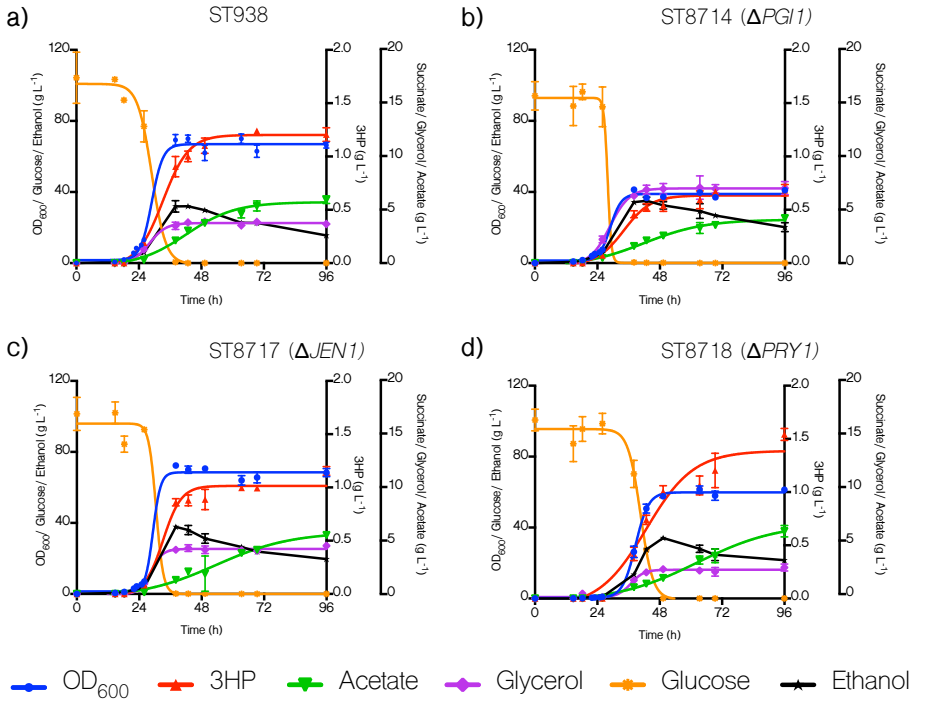


Figure 3. Shake flask fermentations of engineered ST938 strains carrying predicted beneficial gene deletions. a) the fermentation profile for the parental strain 3HP producing strain ST938. b) the PGI1 deletion strain, ST8714. c) the JEN1 deletion strain, ST8717. d) the PRY1 deletion strain, ST8718. Each strain was fermented for 96 hrs in defined mineral media containing 100 g L^{-1} initial glucose concentrations. Each cultivation was performed in triplicate, with the error bars showing the standard deviation for each measurement.

is therefore capable of regenerating the $NADP^+$ required for glucose oxidation through the PP pathway, and can couple glucose assimilation to product formation. Increased flux through the PP pathway should reduce the formation of ethanol, and increase the flux to product formation (Heux et al., 2008). We successfully created a *PGI1* knock-out in ST938 (ST8714), and as expected this strain had a lower growth rate compared to the parental strain (0.198 h^{-1} ± 0.004 *vs.* 0.246 h^{-1} ± 0.004), though this was considerably higher than that reported in the bacterial transhydrogenase overexpression strain constructed by Fiaux

et al., (2003). Total biomass accumulation was much lower for this strain compared to the parental strain. Under shake flask fermentations containing 100 g L^{-1} glucose, ST8714 could only reach a final optical density of approximately 60% of ST938 (OD_{600} 41 ± 1.73 *vs.* 66 ± 1.53). 3HP production was also greatly reduced in ST8714, producing only ~50% as much as ST938 (Table 1). Glycerol production on the other hand was greatly increased, with ~100% more glycerol excreted by the end of the fermentation, indicating that this strain may have higher levels of NADH than ST938, that is then reoxidised through glycerol formation (Ansell et al., 1997; Nissen et al., 2000). In spite of such low 3HP titres, this strategy may show promise for future engineering attempts as ST8714 was able to fully utilize the glucose in the media with a decreased amount of carbon flux diverted into biomass accumulation. Further work is needed to better characterize this strain, as such a shift in carbon assimilation will cause wide reaching effects on regulatory mechanisms and metabolic pathways. The higher glycerol production indicates that ST8714 may have a redox imbalance, or subject to other stressors, and identification strategies to relieve these may boost 3HP titres and productivity.

PRY1 was highly upregulated in ST938 *vs.* the WT strain in fed batch bioreactor fermentations. Pry1p has been identified as a secreted protein that binds acetylated sterols, and facilitates their solubilisation and transport out of the cell (Choudhary and Schneiter, 2012). Mutants deficient in *PRY1* have also been shown to be hypersensitive to eugenol, suggesting that Pry1p can also bind small hydrophobic molecules, helping in their detoxification (Choudhary and Schneiter, 2012). However, in a *S. cerevisiae* isolate tolerant to the organic solvent isooctane, *PRY1* was significantly downregulated (Miura et al., 2000). For these reasons, *PRY1* deleted to investigate its role in 3HP production. Our *PRY1* knockout strain, ST8718, showed a slightly reduced growth rate and biomass accumulation compared to the parental strain and had a 12 hour longer lag phase than any of the other tested strains (Figure 3d). The levels of secreted products were similar between

ST938 and ST8718, but after 96 hours ST8718 had produced ~27% more 3HP. The cause for the increase in 3HP titres is not immediately evident from the limited literature on Pry1p function, but it appeared to continue producing ethanol into the ethanol utilisation stage of the fermentation. Due to the relatively little change in secreted products, it may be assumed that Pry1p does not directly affect the central carbon metabolism, and that its role in either acetylated sterol transport or stress response pathways influence 3HP production. It is possible that such a deletion may cause the strain to be less resistant to stresses faced during fermentation, which may limit its beneficial impact when fermentations are scaled up and increasing stresses are faced by the cells. We tested ST8718 against ST938 for their abilities to withstand different stress factors. A spot test revealed that ST8718 had much less robust growth than ST938 on the plates even under control conditions. While ST938 growth was relatively unaffected by an incubation temperature of 37 ˚C or the addition of 2 mM H_2O_2, ST8718 was more sensitive to both of these stresses (Figure S4). Interestingly however, ST8718 appeared slightly less affected than ST938 when subjected to stress from high (30 g L^{-1}) 3HP concentrations.

4 Conclusions and future perspectives

Through this study we have provided data on the transcriptional and flux level activities of a yeast strain capable of high-level production of 3HP (ST938). From the transcriptional analysis we discovered that much of the central carbon metabolism was upregulated in ST938. The relative carbon flux through the PP pathway was increased in ST938, while it was decreased in the TCA cycle, with more pyruvate flux being successfully diverted into oxaloacetate.

We showed that the 3HP pathway was sufficient to enable the growth of a *PGI1* deletion strain on glucose. Although higher production of 3HP was not achieved in this strain, it provides an example of how it may be able to couple flux through the PP pathway to 3HP

production, and further work on this may lead to a *PGI1* deletion strain that can be an efficient host. The monocarboxylic acid transporter *JEN1* is perhaps a likely candidate for 3HP transport owing to the similar chemical structures of these compounds. We tested the effect of Jen1p on 3HP production and found that it is unlikely to be involved in the export of 3HP. We also tested the impact of *PRY1* deletion on 3HP titres and this was found to improve final 3HP titres by ~27%, although the exact causes for this improvement are unclear. Pry1p has been implicated in stress responses and may as such have an indirect impact on 3HP biosynthesis.

Acknowledgement

This work was funded by the Novo Nordisk Foundation.

The authors declare no financial or commercial conflict of interest.

Supplementary Information

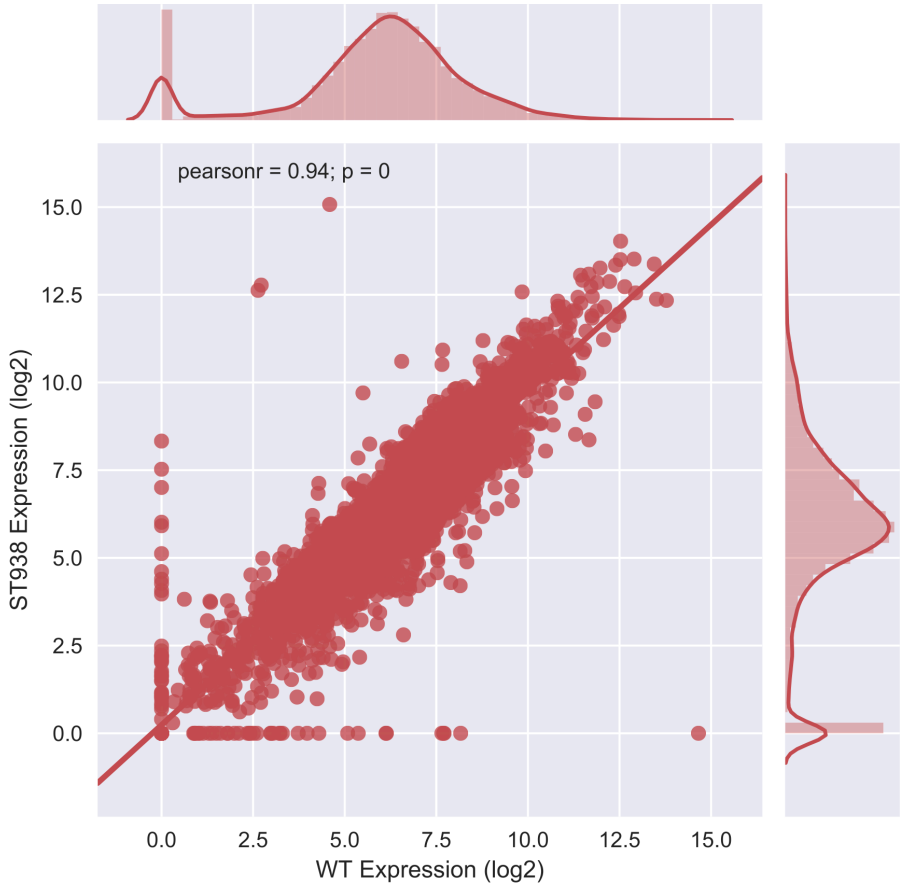


Figure S1. Correlation analysis of the WT vs. ST938 transcriptomes. Values shown are the log_2 transformed gene expression values for each strain. Both strains were grown in 1 L bioreactors, with samples taken for RNAseq analysis once the strains had reached steady state.

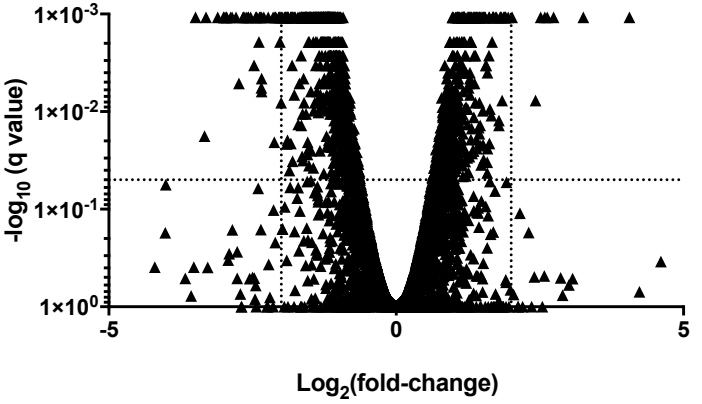


Figure S2. Volcano plot of the RNAseq data for ST938 and the WT strain. The vertical dotted lines represent where the log_2 fold change of gene expression is 2, or -2. Above the horizontal dotted line is where the results fall below the false discovery rate (q-value) significance threshold of 0.05.

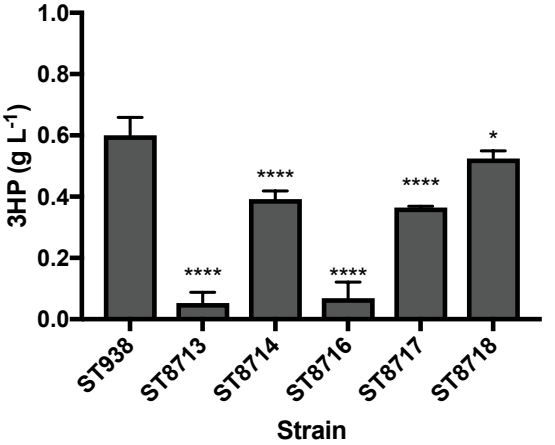


Figure S3. Impact of computationally predicted gene knock-out strategies on the in vivo production of 3HP. Five gene knock-out strategies were implemented into the parental 3HP producer ST938. ST8713 - ΔPSD1, ST8714 - ΔPGI1, ST8716 - ΔOSM1, ST8717 - ΔJEN1, ST8718 - ΔPRY1. Each strain was cultivated in 96-deep well plates in quadruplicate in defined mineral media containing 50 g L^{-1}.

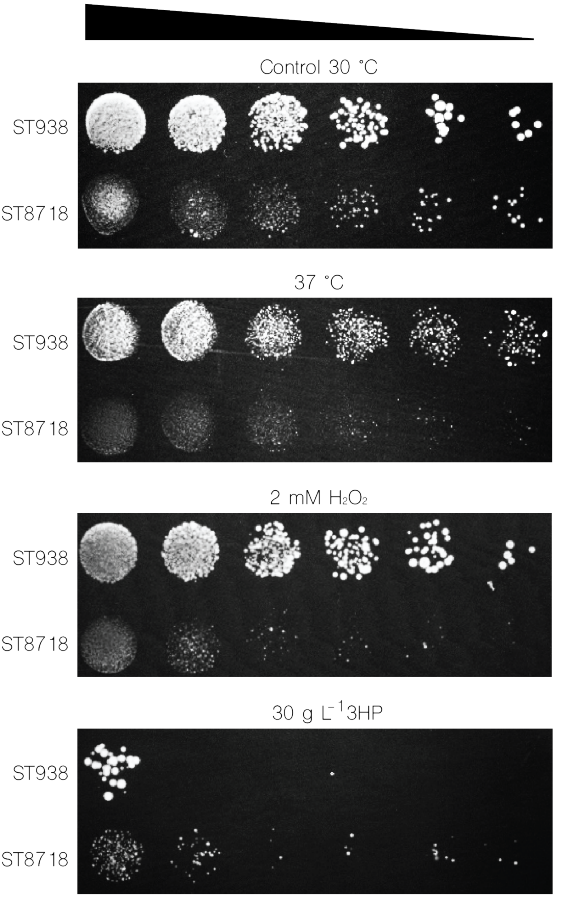


Figure S4. Stress tolerance of a PRY1 deletion in ST938. The parental 3HP producing strain ST938 was compared to the PRY1 deletion strain ST8718 for the inhibition of growth due to different stressors. Both strains were serially diluted (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}) from samples set to the same optical density (OD_{600} 6). The strains were spotted for growth on; YPD at 30 ˚C (the control plate), YPD at 37 ˚C, YPD supplemented with H_2O_2 to a final concentration of 2mM at 30 ˚C, and YPD supplemented with 3-hydroxypropionic acid (3HP) to a final concentration of 30 g L^{-1}.

Table S1. Primers used in this study

Name	Sequence (5' -> 3')
PSD1_KO_fw	GTAATCACATATCACATGCAGGGTA
PSD1_KO_rv	TCATGATGAATAATGCAAAACTCAAGAA
PGI1_KO_fw	GTGTGGGTGTATTGGATTATAGGAA
PGI1_KO_rv	TCAGGTATACTGGAGGCTTCATGAGTTA
NDP3_KO_fw	ACATTCTTGACATCTCAGCTTCTCT
NDP3_KO_rv	TCATTGGTAGATGTTCTTCTAGATTGGC
FRDm_KO_fw	TGAAGTAAAATAGTGATCGGGACTC
FRDm _KO_rv	TCAGTCATTCAGTTTGTTAGAAGAGGGA
L_LACt2r_KO_fw	TATTCTCCACCAGGAAGTTAGTTTG
L_LACt2r_KO_rv	TCAATTGTCTTATAGTGGCAAAAGCAAG
PRY1_KO_fw	ACCCATACCATATATAAACCCACCT
PRY1_KO_rv	TCAATTAAATTGAAATTGTCGAGTGTCG
ScAAT2_U1_fw	AGTGCAGGUAAAACAATGTCTGCCACTCTGTTCA
ScAAT2_U1_rv	CGTGCGAUTTACAATTTAGCTTCAATAGTATAG
PTEF1_fw	ACCTGCACUTTGTAATTAAAACTTAG
PTEF1_rv	CACGCGAUGCACACACCATAGCTTC
TcPAND_U1_fw	AGTGCAGGUAAAACAATGCCAGCTACTGGTG
TcPAND_U1_rv	CGTGCGAUTCACAAATCGGAACCCAATC
PPGK1_rv	ATGACAGAUTTGTTTTATATTTGTTG
EcYdfg_U2_fw	ATCTGTCAUAAAACAATGATCGTTTTAGTAACTGGAG
EcYdfg_U2_rv	CACGCGAUTTACTGACGGTGGACATTC
BcBAPAT_U1_fw	AGTGCAGGUGCATGGTACCAAAACAATG
BcBAPAT_U1_rv	CGTGCGAUATGAGGCCCAGGTCGAC
ScPYC1_U1_fw	AGTGCAGGUAAAACAATGTCGCAAAGAAAATTCG
ScPYC1_U1_rv	CGTGCGAUTCATGCCTTAGTTTCAACAG
ScPYC2_U2_fw	ATCTGTCAUAAAACAATGAGCAGTAGCAAGAAATTG
ScPYC2_U2_rv	CACGCGAUTTACTTTTTTTGGGATGGG
ADH1_test_fw	GAAATTCGCTTATTTAGAAGTGTC
CYC1_test_rv	CTCCTTCCTTTTCGGTTAGAG
TADH1_towards out	GTTGACACTTCTAAATAAGCGAATTTC
X-4-up-out-sq	CTCACAAAGGGACGAATCCT
XI-1-up-sq	CTTAATGGGTAGTGCTTGACACG
XII-1-up-out-sq	CTGGCAAGAGAACCACCAAT

Table S2. Strains used in this study

Name	ID	Description	Source
WT CEN.PK	ST001	CEN.PK WT strain without suxotrophies.	Petter Kötter
ST938	ST938	Strain shown to produce high levels of 3HP through the β-alanine pathway of 3HP biosynthesis	(Borodina et al., 2015)
ST938Δ*PSD1*	ST8713	Deletion of *PSD1* in strain ST938. *PSD1* encodes a phosphatidylserine decarboxylase localised to the mitochondrial inner membrane.	This study
ST938Δ*PGI1*	ST8714	Deletion of *PGI1* in strain ST938. *PGI1* encodes phosphoglucose isomerase that converts glucose-6-phosphate to fructose-6-phosphate in early glycolysis.	This study
ST938Δ*OSM1*	ST8715	Deletion of *OSM1* in strain ST938. *OSM1* encodes fumurate reductase, reducing fumurate to succinate. Located to both the cytosol and to the endoplasmic reticulum.	This study
ST938Δ*JEN1*	ST8717	Deletion of *JEN1* in strain ST938. *JEN1* encodes a monocarboxylate transporter responsible for lactate, pyruvate, and acetate.	This study
ST938Δ*PRY1*	ST8718	Deletion of *PRY1* in strain ST938. *PRY1* encodes a secreted protein involved in the transport of acetylated sterols.	This study

Table S3. Calculated central carbon fluxes

	WT		ST938	
Reaction	**Flux ($mmol^{-1}$ cDW^{-1} h^{-1})**	**StdErr**	**Flux ($mmol^{-1}$ cDW^{-1} h^{-1})**	**StdErr**
GLC + ATP -> G6P	0.463	0.081	0.659	0.037
G6P -> P5P + 2*NADPH +CO2	0.154	0.038	0.244	0.034
G6P -> F6P	0.232	0.062	0.316	0.045
F6P + ATP -> 2T3P	0.321	0.086	0.463	0.066
2P5P -> S7P + T3P	0.05	0.013	0.08	0.012
P5P + E4P -> F6P + T3P	0.039	0.013	0.067	0.013
S7P + T3P -> E4P + F6P	0.05	0.013	0.08	0.012

T3P -> Ser + NADH	0.021	0.015	0.023	0.016
Ser + NADH -> Gly + C1	0.029	0.023	0.022	0.018
Gly + C1 -> Ser + NADH	0.016	0.012	0.008	0.006
C1 + CO2 + NADH -> Gly	0	0.001	0	0.001
T3P -> Pep + ATP + NADH	0.646	0.184	0.954	0.145
PEP -> cytPYR + ATP	0.671	0.199	1.034	0.163
mitPYR -> mitAcCoA + CO2 + NADH	0.316	0.154	0.417	0.133
mitOAA + mitAcCoA -> CIT	0.309	0.154	0.41	0.133
CIT -> OGA + CO2 + NADH	0.309	0.154	0.41	0.133
OGA -> SUC + CO2 + 0.5ATP + NADH	0.27	0.129	0.353	0.111
SUC -> FUM + NADH	0.27	0.129	0.353	0.111
MAL -> mitOAA + NADH	0.248	0.12	0.317	0.099
FUM -> MAL	0.27	0.129	0.353	0.111
MAL -> mitPYR + CO2 + NADPH	0.023	0.012	0.036	0.019
cytOAA + ATP -> PEP + CO2	0.049	0.017	0.106	0.023
cytPYR + CO2 + ATP -> cytOAA	0.164	0.067	0.265	0.067
acetate + 2ATP -> cytAcCoA	0.058	0.056	0.068	0.057
Acetaldehyde -> acetate + NADPH	0.118	0.06	0.104	0.072
Acetaldehyde + NADH -> ethanol	0	0.001	0	0.001
T3P + NADH -> glycerol	0	0.001	0	0.001
cytOAA -> mitOAA	0.099	0.047	0.163	0.05
mitOAA -> cytOAAc	0.038	0.018	0.07	0.022
cytPYR -> mitPYR	0.381	0.153	0.491	0.127
cytPYR -> Acetaldehyde + CO2	0.118	0.06	0.104	0.072
O2 + 2NADH -> 2P/O x ATP	1.038	0.423	1.411	0.348
cytAcCoA +CO2 + 2 NADPH + ATP -> 3HP	0	0.001	0.164	0.029
Biomass	0.041	0.001	0.044	0.01

Table S4. Results from RNAseq. Transcripts which displayed significance (q < 0.05), and had the highest Log_2 fold changes (> 2) are displayed in the table.

Gene Name	Gene ID	Log2 Fold Change (ST938/WT)	q-value
APC4	YDR118W-A	-3.49018	0.00108252
ECM23	YPL021W	-3.33541	0.0178557
CRS5	YOR031W	-3.30663	0.00108252
BOP2	YLR267W	-3.10454	0.00108252
YMR244W	YMR244W	-3.00294	0.00108252
YIL171W-A	YIL171W-A	-2.99842	0.00108252
SRL4	YPL033C	-2.97583	0.00108252
CUP2	YGL166W	-2.9576	0.00108252
SPO20	YMR017W	-2.94711	0.00108252
URA3	YEL021W	-2.87919	0.00108252
YNR064C	YNR064C	-2.85443	0.00108252
SPG1	YGR236C	-2.78502	0.00108252
YMR320W	YMR320W	-2.73402	0.00513697
YNR063W	YNR063W	-2.66495	0.00108252
PHO89	YBR296C	-2.66128	0.00108252
SPO74	YGL170C	-2.63423	0.00108252
YPT53	YNL093W	-2.59801	0.00108252
YOR019W	YOR019W	-2.5753	0.00108252
YGR051C	YGR051C	-2.5702	0.00108252
BDF2	YDL070W	-2.53291	0.00108252
SPG5	YMR191W	-2.47289	0.00336367
ADY2	YCR010C	-2.44088	0.00108252
YHR022C	YHR022C	-2.43229	0.00108252
YMR018W	YMR018W	-2.4242	0.00108252
RGI2	YIL057C	-2.3857	0.00192982
DSF1	YEL070W	-2.3777	0.00108252
YNL033W	YNL033W	-2.36218	0.00108252
UTR5	YEL035C	-2.34454	0.00570254
ECM12	YHR021W-A	-2.34181	0.00622172

Table S4. Results from RNAseq. Transcripts which displayed significance (q < 0.05), and had the highest Log2 fold changes (> 2) are displayed in the table.

NDJ1	YOL104C	-2.34029	0.00459692
YET2	YMR040W	-2.33981	0.00108252
YGR050C	YGR050C	-2.30282	0.00108252
NGL3	YML118W	-2.22753	0.00108252
CSM4	YPL200W	-2.1432	0.00108252
YOR214C	YOR214C	-2.12186	0.020639
CDC31	YOR257W	-2.1123	0.00108252
ALT2	YDR111C	-2.101	0.00108252
UFE1	YOR075W	-2.09601	0.00108252
YER184C	YER184C	-2.04174	0.00108252
DAN1	YJR150C	-2.03683	0.00108252
YER187W	YER187W	-2.02159	0.00192982
RNR3	YIL066C	-2.021	0.00108252
BAP2	YBR068C	-2.01647	0.00108252
NCE101	YJL205C	-2.00146	0.00810498
PRY1	YJL079C	2.019	0.00108252
DAL5	YJR152W	2.42409	0.00769654
MCH5	YOR306C	2.51504	0.00108252
SCW4	YGR279C	2.58759	0.00108252
PYC1	YGL062W	2.62953	0.00108252
PUT1	YLR142W	2.7411	0.00108252
PUT2	YHR037W	3.25557	0.00108252

References

Abbott, D.A., Zelle, R.M., Pronk, J.T., and van Maris, A.J.A. (2009). Metabolic engineering of Saccharomyces cerevisiae for production of carboxylic acids: current status and challenges. FEMS Yeast Res. *9*, 1123–1136.

Aguilera, A. (1986). Deletion of the phosphoglucose isomerase structural gene makes growth and sporulation glucose dependent in Saccharomyces cerevisiae. Mol. Gen. Genet. *204*, 310–316.

Aguilera, A., and Zimmermann, F.K. (1986). Isolation and molecular analysis of the phosphoglucose isomerase structural gene of Saccharomyces cerevisiae. Mol. Gen. Genet. *202*, 83–89.

Akita, O., Nishimori, C., Shimamoto, T., Fujii, T., and Iefuji, H. (2000). Transport of pyruvate in Saccharomyces cerevisiae and cloning of the gene encoded pyruvate permease. Biosci. Biotechnol. Biochem. *64*, 980–984.

Alber, B., Olinger, M., Rieder, A., Kockelkorn, D., Jobst, B., Hugler, M., and Fuchs, G. (2006). Malonyl-coenzyme A reductase in the modified 3-hydroxypropionate cycle for autotrophic carbon fixation in archaeal Metallosphaera and Sulfolobus spp. J. Bacteriol. *188*, 8551–8559.

Ansell, R., Granath, K., Hohmann, S., Thevelein, J.M., and Adler, L. (1997). The two isoenzymes for yeast NAD+-dependent glycerol 3-phosphate dehydrogenase encoded by GPD1 and GPD2 have distinct roles in osmoadaptation and redox regulation. EMBO J. *16*, 2179–2187.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B *57*, 289–300.

Boles, E., Lehnert, W., and Zimmermann, F.K. (1993). The role of the NAD-dependent glutamate dehydrogenase in restoring growth on glucose of a Saccharomyces cerevisiae phosphoglucose isomerase mutant. Eur. J. Biochem. *217*, 469–477.

Borodina, I., Kildegaard, K.R., Jensen, N.B., Blicher, T.H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M.J., Rosenstand, I., et al. (2015). Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. Metab. Eng. *27*, 57–64.

Burgard, A.P., Pharkya, P., and Maranas, C.D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

Cardoso, J.G.R., Jensen, K., Lieven, C., Lærke Hansen, A.S., Galkina, S., Beber, M., Özdemir, E., Herrgård, M.J., Redestig, H., and Sonnenschein, N. (2018). Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories. ACS Synth. Biol. *7*, 1163–1166.

Casal, M., Paiva, S., Andrade, R.P., Gancedo, C., and Leao, C. (1999). The lactate-proton symport of Saccharomyces cerevisiae is encoded by JEN1. J. Bacteriol. *181*, 2620–2623.

Casal, M., Paiva, S., Queirós, O., and Soares-Silva, I. (2008). Transport of carboxylic acids in yeasts. FEMS Microbiol. Rev. *32*, 974–994.

Chen, Y., Daviet, L., Schalk, M., Siewers, V., and Nielsen, J. (2013). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. Metab. Eng. *15*, 48–54.

Choudhary, V., and Schneiter, R. (2012). Pathogen-Related Yeast (PRY) proteins and members of the CAP superfamily are secreted sterol-binding proteins. Proc. Natl. Acad. Sci. U. S. A. *109*, 16882–16887.

Clancey, C.J., Chang, S.C., and Dowhan, W. (1993). Cloning of a gene (PSD1) encoding phosphatidylserine decarboxylase from Saccharomyces cerevisiae by complementation of an Escherichia coli mutant. J. Biol. Chem. *268*, 24580–24590.

Dürrschmid, K., Reischer, H., Schmidt-Heck, W., Hrebicek, T., Guthke, R., Rizzi, A., and Bayer, K. (2008). Monitoring of transcriptome and proteome profiles to investigate the cellular response of E. coli towards recombinant protein expression under defined chemostat conditions. J. Biotechnol. *135*, 34–44.

Fiaux, J., Cakar, Z.P., Sonderegger, M., Wuthrich, K., Szyperski, T., and Sauer, U. (2003). Metabolic-flux profiling of the yeasts Saccharomyces cerevisiae and Pichia stipitis. Eukaryot. Cell *2*, 170–180.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. *2*, 31–34.

Gustavsson, M., and Lee, S.Y. (2016). Prospects of microbial cell factories developed through systems metabolic engineering. Microb. Biotechnol. *9*, 610–617.

Haddadin, F.T., and Harcum, S.W. (2005). Transcriptome profiles for high-cell-density recombinant and wild-

type Escherichia coli. Biotechnol. Bioeng. *90*, 127–153.

Haurie, V., Perrot, M., Mini, T., Jeno, P., Sagliocco, F., and Boucherie, H. (2001). The transcriptional activator Cat8p provides a major contribution to the reprogramming of carbon metabolism during the diauxic shift in Saccharomyces cerevisiae. J. Biol. Chem. *276*, 76–85.

Heux, S., Cadiere, A., and Dequin, S. (2008). Glucose utilization of strains lacking PGI1 and expressing a transhydrogenase suggests differences in the pentose phosphate capacity among Saccharomyces cerevisiae strains. FEMS Yeast Res. *8*, 217–224.

Hoffmann, F., and Rinas, U. (2004). Stress Induced by Recombinant Protein Production in Escherichia coli. In Physiological Stress Responses in Bioprocesses, (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 73–92.

Holo, H. (1989). Chloroflexus aurantiacus secretes 3-hydroxypropionate, a possible intermediate in the assimilation of CO2 and acetate. Arch. Microbiol. *151*, 252–256.

Hugler, M., Menendez, C., Schagger, H., and Fuchs, G. (2002). Malonyl-coenzyme A reductase from Chloroflexus aurantiacus, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO(2) fixation. J. Bacteriol. *184*, 2404–2410.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., and Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. *14*, 238–248.

Kildegaard, K.R., Hallstrom, B.M., Blicher, T.H., Sonnenschein, N., Jensen, N.B., Sherstyk, S., Harrison, S.J., Maury, J., Herrgard, M.J., Juncker, A.S., et al. (2014). Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. Metab. Eng. *26*, 57–66.

Kildegaard, K.R., Jensen, N.B., Schneider, K., Czarnotta, E., Ozdemir, E., Klein, T., Maury, J., Ebert, B.E., Christensen, H.B., Chen, Y., et al. (2016). Engineering and systems-level analysis of Saccharomyces cerevisiae for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway. Microb. Cell Fact. *15*, 53.

Kozak, B.U., van Rossum, H.M., Luttik, M.A.H., Akeroyd, M., Benjamin, K.R., Wu, L., de Vries, S., Daran, J.-M., Pronk, J.T., and van Maris, A.J.A. (2014). Engineering acetyl coenzyme A supply: functional expression of a bacterial pyruvate dehydrogenase complex in the cytosol of Saccharomyces cerevisiae. MBio *5*, e01696-14.

Kumar, V., Ashok, S., and Park, S. (2013). Recent advances in biological production of 3-hydroxypropionic acid. Biotechnol. Adv. *31*, 945–961.

Liu, R., Liang, L., Choudhury, A., Bassalo, M.C., Garst, A.D., Tarasava, K., and Gill, R.T. (2018). Iterative genome editing of Escherichia coli for 3-hydroxypropionic acid production. Metab. Eng. *47*, 303–313.

Liu, W., Zhang, B., and Jiang, R. (2017). Improving acetyl-CoA biosynthesis in Saccharomyces cerevisiae via the overexpression of pantothenate kinase and PDH bypass. Biotechnol. Biofuels *10*, 41.

Lv, X., Wang, F., Zhou, P., Ye, L., Xie, W., Xu, H., and Yu, H. (2016). Dual regulation of cytoplasmic and mitochondrial acetyl-CoA utilization for improved isoprene production in Saccharomyces cerevisiae. *7*, 12851.

Makuc, J., Paiva, S., Schauen, M., Kramer, R., Andre, B., Casal, M., Leao, C., and Boles, E. (2001). The putative monocarboxylate permeases of the yeast Saccharomyces cerevisiae do not transport monocarboxylic acids across the plasma membrane. Yeast *18*, 1131–1143.

Van Maris, A., N. Konings, W., P.van Dijken, J., and Pronk, J. (2004). Microbial export of lactic and 3-hydroxypropanoic acid: Implications for industrial fermentation processes.

Miura, S., Zou, W., Ueda, M., and Tanaka, A. (2000). Screening of Genes Involved in Isooctane Tolerance in Saccharomyces cerevisiae by Using mRNA Differential Display. Appl. Environ. Microbiol. *66*, 4883–4889.

Mo, M.L., Palsson, B.O., and Herrgard, M.J. (2009). Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst. Biol. *3*, 37.

Muratsubaki, H., and Enomoto, K. (1998). One of the fumarate reductase isoenzymes from Saccharomyces cerevisiae is encoded by the OSM1 gene. Arch. Biochem. Biophys. *352*, 175–181.

Nielsen, J. (2014). Synthetic Biology for Engineering Acetyl Coenzyme A Metabolism in Yeast. MBio *5*, e02153-14.

Nissen, T.L., Kielland-Brandt, M.C., Nielsen, J., and Villadsen, J. (2000). Optimization of Ethanol Production in Saccharomyces cerevisiae by Metabolic Engineering of the Ammonium Assimilation. Metab. Eng. *2*, 69–77.

Petzold, C.J., Chan, L.J.G., Nhan, M., and Adams, P.D. (2015). Analytics for Metabolic Engineering. Front. Bioeng. Biotechnol. *3*, 135.

Pronk, J.T., Yde Steensma, H., and Van Dijken, J.P. (1996). Pyruvate metabolism in Saccharomyces cerevisiae. Yeast *12*, 1607–1633.

Reihl, P., and Stolz, J. (2005). The Monocarboxylate Transporter Homolog Mch5p Catalyzes Riboflavin (Vitamin B2) Uptake in Saccharomyces cerevisiae. J. Biol. Chem. *280*, 39809–39817.

van Rossum, H.M., Kozak, B.U., Pronk, J.T., and van Maris, A.J.A. (2016). Engineering cytosolic acetyl-coenzyme A supply in Saccharomyces cerevisiae: Pathway stoichiometry, free-energy conservation and redox-cofactor balancing. Metab. Eng. *36*, 99–115.

Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.-. K., and Keasling, J.D. (2007). Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metab Eng *9*.

Shin, J.H., Kim, H.U., Kim, D.I., and Lee, S.Y. (2013). Production of bulk chemicals via novel metabolic pathways in microorganisms. Biotechnol. Adv. *31*, 925–935.

Smallbone, K., Messiha, H.L., Carroll, K.M., Winder, C.L., Malys, N., Dunn, W.B., Murabito, E., Swainston, N., Dada, J.O., Khan, F., et al. (2013). A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. Febs Lett. *587*, 2832–2841.

Sobolov, M., and Smiley, K.L. (1960). Metabolism of glycerol by an acrolein-forming lactobacillus. J. Bacteriol. *79*, 261–266.

Song, C.W., Kim, J.W., Cho, I.J., and Lee, S.Y. (2016). Metabolic Engineering of Escherichia coli for the Production of 3-Hydroxypropionic Acid and Malonic Acid through beta-Alanine Route. ACS Synth. Biol. *5*, 1256–1263.

Tabita, F.R. (2009). The hydroxypropionate pathway of CO2 fixation: Fait accompli. Proc. Natl. Acad. Sci. U. S. A. *106*, 21015–21016.

Talarico, T.L., and Dobrogosz, W.J. (1990). Purification and Characterization of Glycerol Dehydratase from Lactobacillus reuteri. Appl. Environ. Microbiol. *56*, 1195–1197.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562–578.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. *11*, R14.

Zamboni, N., Fischer, E., and Sauer, U. (2005). FiatFlux--a software for metabolic flux analysis from 13C-glucose experiments. BMC Bioinformatics *6*, 209.

Zarzycki, J., Brecht, V., Muller, M., and Fuchs, G. (2009). Identifying the missing steps of the autotrophic 3-hydroxypropionate CO2 fixation cycle in Chloroflexus aurantiacus. Proc. Natl. Acad. Sci. U. S. A. *106*, 21317–21322.

5 Conclusions and Future Perspectives

The work that has been presented in this thesis has aimed to contribute towards metabolic engineering efforts in *Saccharomyces cerevisiae*. I have presented three research studies into the engineering of *S. cerevisiae* for the production of hydroxy acids. In the following section I will attempt to provide a summary of these works, along with a view for the future development of the yeast cell factory.

New tools for genetic engineering have been presented in Chapter 2, with the EasyClone-MarkerFree methodology. This study provides a CRISPR-Cas9 toolkit for the chromosomal integration of genes without the requirement marker insertion, and also aims to advance standardisation within the field by using the standard EasyClone integration sites and cloning methodology. Genomic integration is a prerequisite for reproducibility as plasmid based systems are inherently noisy (Jensen et al., 2014). The use of defined, characterised, and standardised integration sites is also a prerequisite, as different sites of integration will show different expression profiles (Jessop-Fabre et al., 2016). It is hoped that this toolkit offers a user friendly method for genome engineering in yeast and will be used by the yeast community for engineering strains. In this paper, we used our methodology to test three different cytosolic acetyl-CoA supply strategies for improving 3-hydroxypropionic acid (3HP) production in yeast. From this we discovered that the expression of a bacterial pyruvate dehydrogenase complex was able to increase 3HP titres by up to 95%. This result is a promising first step into the further engineering of this pathway to 3HP biosynthesis. Further study into this pathway is needed to resolve how this strategy helps to boost 3HP, and perhaps offer insights into how this process may be further optimised. Research into acetyl-CoA supply is also relevant for the production of many other chemicals, such as fatty acid ethyl esters that use acetyl-CoA as a building block for their production.

Chapter 3 presents research into the production of hydroxy acids in a Crabtree-negative strain of *S. cerevisiae*, the TAM strain (van Maris et al., 2004). This study uses metabolic flux analysis to uncover the metabolic changes that occur in the host strain when heterologous pathways for hydroxy acid production are inserted. We hope is that this data can help guide future engineering attempts in the TAM strain background. Currently, not much research has been done into the use of this host strain for 3HP production, but our results show that it may be a promising line of further inquiry. Learning from how the introduction of lactate production can restore some of the growth defects of the TAM strain may lead to novel strategies for 3HP production, and increase the productivity of this host beyond that of the WT background. At the centre of this is likely to be the detailed analysis and engineering of the redox cycles, that are of unquestionable importance to product formation strategies in yeast.

In the final chapter of research, metabolic flux and transcriptomic analysis is used to characterise another 3HP producing strain, ST938. We present the results of this analysis, highlighting where there may be issues that are hindering higher 3HP titres. We used computational prediction algorithms to generate gene deletion strategies from the metabolic flux analysis that would be capable of improving the performance of ST938. It was however, a prediction gained from the transcriptomic data that yielded a strain with improved production capabilities. The deletion of *PRY1* increased 3HP titres by 27%, but the role of this gene is not well characterised. Implicated in the transport of acetylated sterols out of the cell, questions remain as to the mechanisms behind the improved 3HP titres. Further study into the effects of this gene deletion in the production of other compounds may help to determine the role of this gene. Further analysis into the effects that this deletion has on the fitness of the strain is also desirable, as such a deletion may prove deleterious to the performance of the strain under stresses faced in large scale fermentations. We also created a *PGI1* deletion strain. In a WT strain, deletion of *PGI1* prevents growth on glucose (Aguilera, 1986). However, the recycling of NADPH by the

3HP pathway allowed for the strain to still growth on glucose as the sole carbon source after deletion of *PGI1*. Although this strain did not produce high 3HP titres, it does open the door to future strategies to engineer the pentose phosphate pathway for increasing flux to 3HP. Finally, we deleted a monocarboxylate transporter gene, *JEN1*, and showed that this does not decrease the production of 3HP, providing evidence that this transporter is not likely to be involved in 3HP secretion. Inefficient product transport can negatively impact on the productivity of the host strain, and therefore the elucidation of native and heterologous transporters can help to improve the efficiency of biochemical production (Pérez-García and Wendisch, 2018; Zelle et al., 2008). Hence, the further study of 3HP transport mechanisms in yeast is a promising area of future research to increase the performance of ST938.

The yeast, *S. cerevisiae*, has long been used in industrial processes for the production of ethanol, both in the beverage industry, and also for the production of bioethanol (Mohd Azhar et al., 2017). It has become a staple in the biopharmaceutical industry, too, for example in the production of recombinant proteins (Ferrer-Miralles et al., 2009; Nielsen, 2013). In order to fully realise the potential of yeast as a production organism, it is necessary to gain a full and detailed understanding of its metabolism, and physiology. To gain this understanding, tools must be concurrently developed to engineer, analyse, and model the host. Great steps are being taken in all areas, with genetic tools allowing for facile gene insertions, deletions, and manipulations (Jakočiūnas et al., 2015; Kim et al., 2016; Mans et al., 2015; Reider Apel et al., 2017; Sander and Joung, 2014; Vanegas et al., 2017). Techniques are rapidly being developed that allow for fine control over pathway expression, and complex regulatory behaviour (Ajikumar et al., 2010; Dueber et al., 2009; Jones et al., 2015; Lee et al., 2013).

-Omic analyses are now offering metabolic engineers an unprecedented amount of information over their strains (Alonso-Gutierrez et al., 2015; Chen and Nielsen, 2013;

Giaever et al., 2002; Kildegaard et al., 2016; van Maris et al., 2004; Petzold et al., 2015). Metabolic models are constantly being refined, with data from omic technologies helping to inform and improve those models (Chandrasekaran and Price, 2010; Fong et al., 2005; Heavner et al., 2013; Hucka et al., 2003; MacDonald et al., 2011). Steps too are being taken to minimise the amount of scientist man-hours that are currently spent on strain engineering by automating strain design and construction. Automation can take over the tedious, time-consuming, and error-prone tasks in routine metabolic engineering practices (Carbonell et al., 2018; Shih et al., 2015; Si et al., 2017). Scientists have then more time to focus on generating novel hypotheses, while at the same time producing more reproducible data. Many of these technologies, while progressing rapidly, are still in nascent form. Research must continue to be applied in these areas, and where possible are adapted to suit high throughput techniques.

With the rapidly decreasing costs of DNA synthesis, cloning may soon not be a time-consuming endeavour for laboratory scientists. Focus will then be shifted to the creation of ever larger libraries based on computational design strategies. Key to this will be the development of design of experiment tools that will be able to predict strategies capable of efficiently probing the possible phenotypic space. Coupled to this will be the need for ever improved methods for high throughput analysis of the created strains, both in the physical analysis, and in the computational methods needed to sift through the vast arrays of data generated.

Undoubtedly, ever larger quantities of data will be produced, and so computational methods will need to keep pace, to integrate new data, and importantly, to draw lessons from them. The field of machine learning will need to be seamlessly woven into the fabric of metabolic engineering to predict new strain designs from generated data sets (Costello and Martin, 2018; Cuperlovic-Culf, 2018). Much work is still needed to increase the

modularity, orthogonality, and predictability of biological circuits so that it becomes as trivial to design a robust gene circuit as it is to build an electrical one.

5.1 References

Aguilera, A. (1986). Deletion of the phosphoglucose isomerase structural gene makes growth and sporulation glucose dependent in Saccharomyces cerevisiae. Mol. Gen. Genet. *204*, 310–316.

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E.J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., and Stephanopoulos, G. (2010). Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in Escherichia coli Science (80-.). *330*, 70 LP-74.

Alonso-Gutierrez, J., Kim, E.-. M., Batth, T.S., Cho, N., Hu, Q., and Chan, L.J.G. (2015). Principal component analysis of proteomics (PCAP) as a tool to direct metabolic engineering. Metab Eng *28*.

Carbonell, P., Jervis, A.J., Robinson, C.J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K.A., Currin, A., Rattray, N.J.W., et al. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. Commun. Biol. *1*, 66.

Chandrasekaran, S., and Price, N.D. (2010). Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis. Proc Natl Acad Sci USA *107*.

Chen, Y., and Nielsen, J. (2013). Advances in metabolic pathway and strain engineering paving the way for sustainable production of chemical building blocks. Curr. Opin. Biotechnol. *24*, 965–972.

Costello, Z., and Martin, H.G. (2018). A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. Npj Syst. Biol. Appl. *4*, 19.

Cuperlovic-Culf, M. (2018). Machine Learning Methods for Analysis of Metabolic Data and Metabolic Pathway Modeling. Metabolites *8*.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A. V, Prather, K.L.J., and Keasling, J.D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. *27*, 753–759.

Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J.L., Vazquez, E., and Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. Microb. Cell Fact. *8*, 17.

Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., and Palsson, B.O. (2005). In silico design and adaptive evolution of Escherichia coli for production of lactic acid. Biotechnol. Bioeng. *91*,

643–648.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Heavner, B.D., Smallbone, K., Price, N.D., and Walker, L.P. (2013). Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance. Database (Oxford) *2013*.

Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics *19*, 524–531.

Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., and Pedersen, L.E. (2015). Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab Eng *28*.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., and Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. *14*, 238–248.

Jessop-Fabre, M.M., Jakociunas, T., Stovicek, V., Dai, Z., Jensen, M.K., Keasling, J.D., and Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. Biotechnol. J. *11*, 1110–1117.

Jones, J.A., Toparlak, O.D., and Koffas, M.A.G. (2015). Metabolic pathway balancing and its role in the production of biofuels and chemicals. Curr. Opin. Biotechnol. *33*, 52–59.

Kildegaard, K.R., Jensen, N.B., Schneider, K., Czarnotta, E., Ozdemir, E., Klein, T., Maury, J., Ebert, B.E., Christensen, H.B., Chen, Y., et al. (2016). Engineering and systems-level analysis of Saccharomyces cerevisiae for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway. Microb. Cell Fact. *15*, 53.

Kim, S.K., Han, G.H., Seong, W., Kim, H., Kim, S.-. W., and Lee, D.-. H. (2016). CRISPR interference-guided balancing of a biosynthetic mevalonate pathway increases terpenoid production. Metab Eng.

Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J., and Dueber, J.E. (2013). Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Res *41*.

MacDonald, J.T., Barnes, C., Kitney, R.I., Freemont, P.S., and Stan, G.-B. V (2011). Computational design approaches and tools for synthetic biology. Integr. Biol. (Camb). *3*, 97–108.

Mans, R., Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., and Daran-lapujade, P. (2015). CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res *15*.

van Maris, A.J.A., Geertman, J.-M.A., Vermeulen, A., Groothuizen, M.K., Winkler, A.A., Piper, M.D.W., van Dijken, J.P., and Pronk, J.T. (2004). Directed Evolution of Pyruvate Decarboxylase-Negative Saccharomyces cerevisiae, Yielding a C2-Independent, Glucose-Tolerant, and Pyruvate-Hyperproducing Yeast. Appl. Environ. Microbiol. *70*, 159–166.

Mohd Azhar, S.H., Abdulla, R., Jambo, S.A., Marbawi, H., Gansau, J.A., Mohd Faik, A.A., and Rodrigues, K.F. (2017). Yeasts in sustainable bioethanol production: A review. Biochem. Biophys. Reports *10*, 52–61.

Nielsen, J. (2013). Production of biopharmaceutical proteins by yeast: Advances through metabolic engineering. Bioengineered *4*, 207–211.

Pérez-García, F., and Wendisch, V.F. (2018). Transport and metabolic engineering of the cell factory Corynebacterium glutamicum. FEMS Microbiol. Lett. *365*, fny166-fny166.

Petzold, C.J., Chan, L.J.G., Nhan, M., and Adams, P.D. (2015). Analytics for Metabolic Engineering. Front. Bioeng. Biotechnol. *3*, 135.

Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R.A., Tong, G.J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N.J., et al. (2017). A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. Nucleic Acids Res. *45*, 496–508.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotech *32*, 347–355.

Shih, S.C.C., Goyal, G., Kim, P.W., Koutsoubelis, N., Keasling, J.D., Adams, P.D., Hillson, N.J., and Singh, A.K. (2015). A Versatile Microfluidic Device for Automating Synthetic Biology. ACS Synth. Biol. *4*, 1151–1164.

Si, T., Chao, R., Min, Y., Wu, Y., Ren, W., and Zhao, H. (2017). Automated multiplex genome-scale

engineering in yeast. Nat. Commun. *8*, 15187.

Vanegas, K.G., Lehka, B.J., and Mortensen, U.H. (2017). SWITCH: a dynamic CRISPR tool for genome engineering and metabolic pathway control for cell factory construction in Saccharomyces cerevisiae. Microb Cell Fact *16*.

Zelle, R.M., de Hulster, E., van Winden, W.A., de Waard, P., Dijkema, C., Winkler, A.A., Geertman, J.-M.A., van Dijken, J.P., Pronk, J.T., and van Maris, A.J.A. (2008). Malic Acid Production by Saccharomyces cerevisiae: Engineering of Pyruvate Carboxylation, Oxaloacetate Reduction, and Malate Export . Appl. Environ. Microbiol. *74*, 2766–2777.